Quantum Elimination Measurements

Jonathan Crickmore

A thesis presented for the degree of
Doctor of Philosophy

Heriot-Watt University

School of Engineering and Physical Sciences

Institute of Photonics and Quantum Sciences

August 2019

Abstract

If an initial state is prepared from a known set, then the aim of a quantum state elimination measurement is to rule out a subset of the possible initial states. We use semidefinite programming to find either bounds or exact results on the success probabilities of certain elimination measurements. In conjunction we use an analytic approach to find optimal measurements. We obtain optimal measurements for unambiguous elimination in a two-qubit case where each qubit is in one of two possible states. We also show how it might be possible to use our elimination measurements in a QKD protocol. In addition we prove that the best method to eliminate the highest average number of states for sequences of qubits with each qubit in one of two possible states is individual unambiguous measurements. Furthermore we show the method of decomposing a unitary matrix into beamsplitter-like operations found by Reck et al. and apply this to our elimination measurement to realise a way of experimental implementation.

In the final chapter we look at joint measurements and find the optimal probe state that we would use to minimise the uncertainty in our estimation of the sharpness of a measurement between two observables.

Acknowledgements

Firstly I would like to thank my supervisor Prof. Erika Andersson for her expertise and continued advice in the subject field. Besides my supervisor I would like to especially thank Dr. Ittoop Puthoor for his help throughout the project on many different aspects. Further thanks to Dr. Petros Wallden, Dr. Sarah Croke and Prof. Mark Hillery.

I would also like to thank the Scottish Doctoral Training Centre for Condensed Matter Physics (CM-CDT), which provided funds through EPSRC, learning opportunities and a cohort of other students to go through this journey with.

Finally a thanks all my family and friends that have supported me over the last four years with support and guidance.

Research Thesis Submission

Please note this form should be bound into the submitted thesis.

Name:	Jonathan Crickmore		
School:	Engineering and Physical Sciences		
Version: *(i.e. First, Resubmission, Final)*	Final	Degree Sought:	Doctor of Philosophy

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1. The thesis embodies the results of my own work and has been composed by myself
2. Where appropriate, I have made acknowledgement of the work of others
3. The thesis is the correct version for submission and is the same version as any electronic versions submitted*.
4. My thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5. I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.
6. I confirm that the thesis has been verified against plagiarism via an approved plagiarism detection application e.g. Turnitin.

ONLY for submissions including published works

Please note you are only required to complete the Inclusion of Published Works Form (page 2) if your thesis contains published works)

7. Where the thesis contains published outputs under Regulation 6 (9.1.2) or Regulation 43 (9) these are accompanied by a critical review which accurately describes my contribution to the research and, for multi-author outputs, a signed declaration indicating the contribution of each author (complete)
8. Inclusion of published outputs under Regulation 6 (9.1.2) or Regulation 43 (9) shall not constitute plagiarism.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	30/01/2020

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Limited Access	Requested	Yes		No		Approved	Yes		No	
*E-thesis Submitted (**mandatory for final theses**)*										
Received in the SSC by *(name in capitals):*						Date:				

Contents

Chapter 1

Introduction

1.1 Background and Motivation

This section is here to give a little information on the motivation for the work in the thesis and how I came to work on elimination measurements.

The interest in elimination measurements came from looking at the Pusey-Barrett-Rudolph (PBR) theorem [1] (which we shall explain in detail later in the introduction). It uses an elimination measurement to perform a task that isn't obviously possible and with little research done in the area we thought it would be interesting to pursue a variety of elimination measurements ourselves. We took the measurement used by PBR as a basis for our work and expanded from there. In the search for a method to obtain some numerical results I came across semi-definite programming (SDP) (chapter 2), which turned out to be a very useful tool for finding the probabilities of success of certain measurements. With this tool we were able to compare numerical results with analytical methods (chapter 3) to check the performance of our measurements.

Once we found the optimal measurements we looked at methods of implementing our measurement (chapter 4) and investigating more general approaches of optimising the process of going from a unitary operator to an experimental setup.

The final chapter on joint measurements was done in collaboration with an experimental group in Bristol and so the motivation was to solve the theoretical problems they had.

In the introduction I will first go through quantum information and measurements in general using unambiguous state discrimination as an example. This is followed by the PBR measurement where the motivation for our work originated along with some other examples of uses of state elimination measurements in current literature.

1.2 Quantum Information

Quantum information refers simply to the information of the state of a quantum system. Quantum measurements are a crucial component of quantum information as they are a method of accessing the the information of a quantum system. There are some very useful guides for understanding quantum information out there and especially measurements out there. A few ones I found very useful were [2, 3, 4].

1.2.1 Pure and Mixed States

The quantum state is fundamental to quantum mechanics and is represented by a state vector (ket) $|\psi\rangle$. A combination of two states say $|\psi\rangle$ and $|\phi\rangle$ is also a quantum state. A pure state can not be written as a statistical ensemble of other states. A mixture of pure states is known as a mixed state and is often represented by a density matrix ρ,

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \tag{1.1}$$

where p_j is the probability the system is in the pure state $|\psi_j\rangle$. Of course we can just write a pure state $|\psi\rangle$ as $\rho = |\psi\rangle\langle\psi|$ and the calculations will work. In the Bloch sphere picture a pure state will lie on the surface of the sphere and mixed state on the interior. The purity of a state can be seen as how close the state is to the surface or centre of the sphere. The density matrix must satisfy certain requirements. Conversely any operator that satisfies these requirements is a density matrix.

- $Tr(\rho) = 1$ from the fact that

$$Tr(\rho) = Tr\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i = 1. \tag{1.2}$$

- It is hermitian: $\rho = \rho^\dagger$.
- It is positive semi-definite such that $\langle\psi|\rho|\psi\rangle \geq 0$.

Once we have the state we now require a method to investigate it and that is where measurements come in.

1.2.2 Projective Measurements

Many of us are first introduced to quantum measurements with the approach proposed by Von Neumann [5] in 1932. This is also referred to as a projective measurement.
A projective measurement is when the measurement operators M_i are all projectors. A projector *P* is an operator that projects one state to another state, for example

$$P_0 = |0\rangle\langle 0| \tag{1.3}$$

is a projector onto the state $|0\rangle$. The projectors must satisfy the following requirements

- They are Hermitian, $P^\dagger = P$.
- They are positive semi-definite operators, therefore $\langle\psi|P|\psi\rangle \geq 0$ for any state $|\psi\rangle$.
- $P^2 = P$ as $P^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|$ due to the fact $|\langle\psi|\psi\rangle| = 1$.
- The projectors are orthogonal such that $P_i P_{i'} = \delta_{i,i'} P_i$. This comes from the idea you can only get one outcome. This can be proven by using the previous requirements.

The probability of obtaining the outcome related to the projector P_i is given by

$$p(i) = \langle\psi|P_i|\psi\rangle. \tag{1.4}$$

The measurement process is random, we should look at $|\psi\rangle$ as a collection of identically prepared states. If we measure each individual state with the same measurement then we can predict the results and the probabilities with which they occur but we can't predict the individual measurement outcomes. An exception to the final condition that the projectors are orthogonal is if we have an observable with a degenerate eigensystem.

Given outcome i occurred then the post measurement state is

$$|\phi\rangle_i = \frac{P_i|\psi\rangle}{\sqrt{\langle\psi|P_i|\psi\rangle}}, \tag{1.5}$$

or alternatively for a mixed state we have

$$|\phi\rangle_i = \frac{P_i \rho P_i}{Tr(P_i \rho P_i)} \tag{1.6}$$

The final condition of the projectors also implies that there can only be as many projectors as there are dimensions of the measured system as they are orthogonal. It turned out this was not essential and from this generalised measurements were introduced.

1.2.3 Generalised Measurements

For generalised measurements we introduce a set of probability operators M_i, relating to a respective measurement outcome i. These have the conditions that:

- The operators are hermitian: $M_i = M_i^\dagger$.
- They are positive semi-definite operators: $\langle\psi|M_i|\psi\rangle \geq 0$ for all possible $|\psi\rangle$.
- They are complete: $\sum_i M_i = I$.

Each outcome i has a probability of occurring given by $p = Tr[M_i\rho]$, where ρ is the measured state. The set of operators is often referred to as a positive operator-valued measure (POVM). As we now do not require the operators to be orthogonal then it is possible to have more outcomes than dimensions of the system and this is the main advantage of a POVM. A POVM can be realised as a projective measurement on an extended Hilbert space. The extension of the original Hilbert space is described by the Neumark extension [6]. We shall go through this extension in detail in chapter 4 where it is integral to our work.

With a POVM the operators that generate the probabilities are not necessarily the same operators that generate the post-measurement state. If we use A_i to denote the operators that form the post-measurement state such that the normalised post-measurement state is

$$|\phi\rangle_i = \frac{A_i|\psi\rangle}{\sqrt{\langle\psi|A_i^\dagger A_i|\psi\rangle}}. \tag{1.7}$$

The operators A_i are Kraus operators and can be seen as the orthogonal projectors that realise the POVM. These are not unique though and often are not known, therefore the post-measurement state of a POVM is often of no use and the measurement is just used to obtain the probabilities with no regard to the post-measurement state. For a POVM A_i can be just about anything even non-hermitian and the probability operators $M_i = A_i^\dagger A_i$, which is a positive operator by construction. As M_i is positive then $M_i^{1/2}$ exists and the most general construction of A_i is $A_i = U_i M_i^{1/2}$, where U_i is any unitary.

We shall now introduce quantum state elimination and follow on with some examples of the measurement process to help understand the process of formulating measurements.

1.2.4 Quantum State Elimination

The aim of a quantum state elimination (QSE) measurement is to exclude one or many of the possible initial states of a quantum state. This is opposed to quantum state discrimination, where the aim is to determine the state of the system. An obvious example of an elimination measurement is that when you perform any projective measurement in a basis, then we can rule out any state orthogonal to the basis state associated with the outcome we obtained. Alternatively you can have a system with four possible initial states that we shall label $|0\rangle, |1\rangle, |2\rangle$ and $|3\rangle$. If the system was in state $|0\rangle$ then an example of an elimination result would be getting an outcome associated to $|1\rangle, |2\rangle, |3\rangle$ or any combination of the them. Obviously if we eliminate all the other possibilities this becomes state discrimination. Quantum state discrimination has been thoroughly studied but there has been much fewer results with elimination measurements [7, 1, 8, 9, 10].

1.2.5 Finding The Measurement Operators

In the following thesis we regularly use quantum measurement theory to derive and test measurements. The process is very similar in many cases and so I will go through it step by step now to be referred to in the future as reasoning for the process.

There are a set of initial states that are the possible states a system could be found in, and these are given as,

$$|\psi_1\rangle, |\psi_2\rangle ... |\psi_n\rangle. \tag{1.8}$$

There are measurement operators that each correspond to an outcome. In our case of quantum state elimination the projectors onto states that are orthogonal to one or more of the initial states will be important when constructing the measurement operators. When achieving the outcome related to that measurement operator you know there was no possibility the system was initially in the state orthogonal to the state being projected onto. This is due to the fact there is no overlap between a state and its orthogonal state. These orthogonal states are denoted by

$$|\psi_{\bar{1}}\rangle, |\psi_{\bar{2}}\rangle ... |\psi_{\bar{n}}\rangle, \tag{1.9}$$

where $|\psi_{\bar{1}}\rangle$ is orthogonal to $|\psi_1\rangle$ and analogously $|\psi_{\bar{2}}\rangle$ is orthogonal to $|\psi_2\rangle$ and so on for the other states. From these orthogonal states we can form a set of projection operators

$$\begin{aligned} P_{\bar{1}} =& |\psi_{\bar{1}}\rangle\langle\psi_{\bar{1}}| \\ P_{\bar{2}} =& |\psi_{\bar{2}}\rangle\langle\psi_{\bar{2}}| \\ &\vdots \\ P_{\bar{n}} =& |\psi_{\bar{n}}\rangle\langle\psi_{\bar{n}}|. \end{aligned} \tag{1.10}$$

The measurement operators $\Pi_{\bar{1}}$ are proportional to the respective projectors

$$\Pi_{\bar{i}} \propto P_{\bar{i}} = |\psi_{\bar{i}}\rangle\langle\psi_{\bar{i}}|. \tag{1.11}$$

For the completeness equation to hold for the measurement operators we require the sum of the operators to be identity. If for example we have four measurement operators then using the completeness equation we get the condition,

$$\alpha_1|\psi_{\bar{1}}\rangle\langle\psi_{\bar{1}}| + \alpha_2|\psi_{\bar{2}}\rangle\langle\psi_{\bar{2}}| + \alpha_3|\psi_{\bar{3}}\rangle\langle\psi_{\bar{3}}| + \alpha_4|\psi_{\bar{4}}\rangle\langle\psi_{\bar{4}}| + \alpha_f|\psi_f\rangle\langle\psi_f| = I, \tag{1.12}$$

where $\alpha_f|\psi_f\rangle\langle\psi_f|$ represents the failure operator and is required to complete the measurement. It is called the failure operator as if we achieve that outcome we have learnt nothing and so the measurement was deemed a failure. Often we have a case when due to symmetry we believe the weightings on all the terms (excluding the failure operator)

are equal. To calculate the optimal weighting we sum together the projectors and choose the coefficients α_i so that the measurement operators give the highest success probability. The first step is to take the sum of the projectors (not including the failure operator). For example the sum of the four projectors in the two-qubit case can be represented by a 4x4 matrix. The dimensions of the matrix will equal the dimensions of the Hilbert space in which you are measuring.

We require the eigenvalues of the sum of the measurement operators corresponding to a successful outcome to be greater than zero and less than or equal one. This is because the sum of projectors as in (1.12) (without the failure projector) is a measurement operator and for any state $|\psi\rangle$ the probability should lie between zero and one.

To find the eigenvalues, we have to diagonalise the matrix that is proportional to the sum of the measurement operators. Then once this is done you scale the matrix so the largest diagonal element becomes one. This scaling is the weighting applied to the projectors and from that we can derive the measurement operators. We shall look at an example in the next section.

1.2.6 Unambiguous State Discrimination

We will now look at unambiguous state discrimination between two pure states to demonstrate the process of forming the measurement operators. If we take a two dimensional space spanned by the orthogonal basis $\{|0\rangle, |1\rangle\}$, then we can express the two states to be discriminated between as

$$\begin{aligned} |\psi_0\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle, \\ |\psi_1\rangle &= \cos\theta|0\rangle - \sin\theta|1\rangle. \end{aligned} \tag{1.13}$$

The states orthogonal to the states in (1.13) are given by

$$\begin{aligned} |\psi_{\overline{0}}\rangle &= \sin\theta|0\rangle - \cos\theta|1\rangle, \\ |\psi_{\overline{1}}\rangle &= \sin\theta|0\rangle + \cos\theta|1\rangle. \end{aligned} \tag{1.14}$$

Consider then the measurement operators

$$\begin{aligned} \pi_{\overline{0}} &= |\psi_{\overline{0}}\rangle\langle\psi_{\overline{0}}| = (\sin\theta|0\rangle - \cos\theta|1\rangle)(\sin\theta\langle 0| - \cos\theta\langle 1|) \\ &= \begin{pmatrix} \sin^2\theta & -\cos\theta\sin\theta \\ -\cos\theta\sin\theta & \cos^2\theta \end{pmatrix}, \\ \pi_{\overline{1}} &= |\psi_{\overline{1}}\rangle\langle\psi_{\overline{1}}| = (\sin\theta|0\rangle + \cos\theta|1\rangle)(\sin\theta\langle 0| + \cos\theta\langle 1|) \\ &= \begin{pmatrix} \sin^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \cos^2\theta \end{pmatrix}. \end{aligned} \tag{1.15}$$

These will be subnormalised later in the process. If we obtain outcome $\overline{0}$ then we know the state was definitely not $|\psi_0\rangle$ and so must have been $|\psi_1\rangle$. This is because $\langle\psi_0|\pi_0|\psi_{\overline{0}}\rangle = 0$. Similarly if we obtain outcome $\overline{1}$ we know the state was definitely not $|\psi_1\rangle$ and so must have been $|\psi_0\rangle$. In this scenario when there are only two options a successful elimination measurement is also a discrimination measurement. To find the probability of a successful measurement we will now sum up the two measurement operators from (1.15),

$$M = \pi_{\overline{0}} + \pi_{\overline{1}} = \begin{pmatrix} 2\sin^2\theta & 0 \\ 0 & 2\cos^2\theta \end{pmatrix}. \tag{1.16}$$

If $\theta = 45^\circ$ then states $|\psi_0\rangle$ and $|\psi_1\rangle$ are orthogonal and $2\sin^2\theta = 2\cos^2\theta = 1$. Now M becomes the identity and no failure operator is required to complete the measurement. This is expected since it is possible to distinguish between orthogonal states with certainty. For the region $0 \leq \theta < 45^\circ$, we have $2\cos^2\theta > 1 > 2\sin^2\theta$. To scale (1.16) so that the new scaled operator does not have eigenvalues exceeding one we need to scale down $\pi_{\overline{0}} + \pi_{\overline{1}}$ so that the diagonal element $2\cos^2\theta$ becomes one . Therefore the new scaled operator M is as follows,

$$M = \frac{1}{2\cos^2\theta}\begin{pmatrix} 2\sin^2\theta & 0 \\ 0 & 2\cos^2\theta \end{pmatrix} = \tan^2\theta|0\rangle\langle 0| + |1\rangle\langle 1|. \tag{1.17}$$

This has assumed the weighting onto both projectors is the same. The failure operator is then given by

$$\Pi_f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{2\cos^2\theta}\begin{pmatrix} 2\sin^2\theta & 0 \\ 0 & 2\cos^2\theta \end{pmatrix} = \begin{pmatrix} 1-\tan^2\theta & 0 \\ 0 & 0 \end{pmatrix} = (1-\tan^2\theta)|0\rangle\langle 0|. \tag{1.18}$$

The probability of failure can be calculated from equation (**??**).

$$p(f) = \eta_0\langle\psi_0|\Pi_f|\psi_0\rangle + \eta_1\langle\psi_1|\Pi_f|\psi_1\rangle = \text{Tr}[\pi_f(\eta_0\rho_0 + \eta_1\rho_1)], \tag{1.19}$$

where η_0 and η_1 are the a priori probabilities of the states $|\psi_0\rangle$ and $|\psi_1\rangle$ being sent. For $\eta_0 = \eta_1 = 1/2$ then

$$\begin{aligned} p(f) =& \frac{1}{2}(\cos\theta\langle 0| + \sin\theta\langle 1|)((1-\tan^2\theta)|0\rangle\langle 0|)(\cos\theta|0\rangle + \sin\theta|1\rangle) \\ &+ \frac{1}{2}(\cos\theta\langle 0| - \sin\theta\langle 1|)((1-\tan^2\theta)|0\rangle\langle 0|)(\cos\theta|0\rangle - \sin\theta|1\rangle) \\ =& \cos^2\theta(1-\tan^2\theta) = \cos^2\theta - \sin^2\theta = \cos(2\theta). \end{aligned} \tag{1.20}$$

Alternatively you can calculate the success probability straight from the operator M

$$\begin{aligned} p(s) =& \frac{1}{2}(\cos\theta\langle 0| + \sin\theta\langle 1|)M(\cos\theta|0\rangle + \sin\theta|1\rangle) \\ &+ \frac{1}{2}(\cos\theta\langle 0| - \sin\theta\langle 1|)M(\cos\theta|0\rangle - \sin\theta|1\rangle) \\ =& 2\sin^2\theta = 1 - \cos(2\theta) = 1 - p(f). \end{aligned} \tag{1.21}$$

This is in fact the optimal result for an unambiguous measurement for equal prior probabilities and is often referred to as the IDP-limit, named after the results from the papers of Ivanovic [11], Dieks [12] and Peres [13]. Our assumption that the weighting on the operators are equal is valid in the case when $\eta_0 = \eta_1 = 1/2$. A more general solution for different a priori probabilities was later found by Jaeger and Shimony [14]. The optimal result can be found by first taking the measurement operators

$$\begin{aligned} \pi_{\overline{0}} &= |a_0|^2 |\psi_{\overline{0}}\rangle\langle\psi_{\overline{0}}|, \\ \pi_{\overline{1}} &= |a_1|^2 |\psi_{\overline{1}}\rangle\langle\psi_{\overline{1}}|, \end{aligned} \tag{1.22}$$

where $|a_i|^2$ are the weightings for each measurement operator. The probability of each outcome can be calculated as

$$\begin{aligned} \langle\psi_1|\pi_0|\psi_1\rangle &= |a_0|^2 |\langle\psi_1|\psi_{\overline{0}}\rangle|^2 = |a_0|^2 \sin^2(2\theta) = p_1, \\ \langle\psi_0|\pi_1|\psi_0\rangle &= |a_1|^2 |\langle\psi_0|\psi_{\overline{1}}\rangle|^2 = |a_1|^2 \sin^2(2\theta) = p_0, \end{aligned} \tag{1.23}$$

where p_1 and p_0 are the probabilities of identifying $|\psi_1\rangle$ and $|\psi_0\rangle$ respectively. Solving for the weightings a in terms of p and θ we get

$$|a_0|^2 = \frac{p_1}{\sin^2(2\theta)} \qquad \text{and} \qquad |a_1|^2 = \frac{p_0}{\sin^2(2\theta)}. \tag{1.24}$$

Substituting in these weightings to the measurement operators given in (1.22) we get

$$\pi_{\overline{0}} = \frac{p_1}{\sin^2(2\theta)} |\psi_{\overline{0}}\rangle\langle\psi_{\overline{0}}| \qquad \text{and} \qquad \pi_{\overline{1}} = \frac{p_0}{\sin^2(2\theta)} |\psi_{\overline{1}}\rangle\langle\psi_{\overline{1}}|. \tag{1.25}$$

The failure operator is given by $\pi_f = I - \pi_0 + \pi_1$ and this is required to be positive semi-definite. If we form this then by requiring the eigenvalues to be greater or equal to 0 then we get the condition

$$(q_0)(q_1) \geq \cos^2(2\theta). \tag{1.26}$$

where $q_i = 1 - p_i$. The average failure probability can be written as

$$Q = \eta_0 q_0 + \eta 1 q_1, \tag{1.27}$$

then to optimise the measurement we wish to minimise Q. To minimise this we see that we want to saturate the bound in (1.26) as this will give the smallest values for q_i. Therefore we can write $q_0 = \cos^2(2\theta)/q_1$, and rewriting Q in terms of just q_0 we get

$$Q = \eta_0 q_0 + \frac{\eta_1 \cos^2(2\theta)}{q_0}. \tag{1.28}$$

Minimising this we get

$$\begin{aligned} \frac{dQ}{dq_0} &= \eta_0 - \frac{\eta_1 \cos^2(2\theta)}{q_0^2} = 0, \\ q_0 &= \sqrt{\frac{\eta_1}{\eta_0}} \cos(2\theta). \end{aligned} \tag{1.29}$$

Similarly for q_1 we have

$$q_1 = \sqrt{\frac{\eta_0}{\eta_1}}\cos(2\theta). \tag{1.30}$$

We can get the optimal measurement operators by writing $p_i = 1 - q_i$ and substituting into (1.25) for the measurement operators and then into (1.28) to get the minimum failure probability of

$$Q = \frac{1}{2} - \frac{1}{2}\sqrt{(1 - 4\eta_0\eta_1)|\langle\psi_0|\psi_1\rangle|^2}. \tag{1.31}$$

For multiple states Chefles has gone on to show that unambiguous discrimination will only work with linearly independent states [15], yet finding explicit solutions is more complicated when more state are involved [16]. A useful text for looking at state discrimination is by Bergou et al. [17].

1.3 The PBR Measurement

An interesting elimination measurement is that proposed by Pusey, Barrett and Rudolph [1] in a paper to investigate the reality of the wave function. The result of the work is summed up nicely by the authors as:

> *... any model in which a quantum state represents mere information about an underlying physical state of the system, and in which systems that are prepared independently have independent physical states, must make predictions that contradict those of quantum theory.*

This result is obtained by assuming two independently produced quantum states have an overlap in the their probability distributions. These two states are then brought together and a measurement exists that eliminates one of the four possibilities with certainty. Therefore the assumptions that the states' probability distributions overlap is false and therefore the state can't be merely information of the system and must be a physical property instead. This rules out a class of quantum models called ψ-*epistemic* that describe the states as being information instead of ψ-*ontic* which describes the reality based theorems. The measurement used is describe below and is a basis to much of the work in this thesis.

A global measurement is guaranteed to eliminate one of the possible initial states the system may be in. These are given by the following four two-qubit states

$$\begin{aligned} &|0\rangle \otimes |0\rangle, \\ &|0\rangle \otimes |+\rangle, \\ &|+\rangle \otimes |0\rangle, \\ &|+\rangle \otimes |+\rangle, \end{aligned} \tag{1.32}$$

where the first qubit is in either $|0\rangle$ or $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ and similarly for the second qubit. The individual states of the qubits involved are non-orthogonal so an elimination measurement on the individual qubits will only give you success with probability less than one. Success here is defined as being able to eliminate one of the possible initial states given in (1.32). I will go through the measurement as it is a basis for much of the work covered in the thesis and was used as a starting point to investigate similar elimination measurements.

We can form an entangled measurement basis (which we will call the PBR basis), consisting of states that each are orthogonal to one of the states given in (1.32). The basis is given by

$$\begin{aligned}
|\psi_{\overline{00}}\rangle &= \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle), \\
|\psi_{\overline{0+}}\rangle &= \frac{1}{\sqrt{2}}(|1+\rangle + |0-\rangle), \\
|\psi_{\overline{+0}}\rangle &= \frac{1}{\sqrt{2}}(|-0\rangle + |+1\rangle), \\
|\psi_{\overline{++}}\rangle &= \frac{1}{\sqrt{2}}(|-+\rangle + |+-\rangle),
\end{aligned} \tag{1.33}$$

where $|-\rangle = 1/\sqrt{2}(|0\rangle - |1\rangle)$ is the state orthogonal to $|+\rangle$. $|\psi_{\overline{00}}\rangle$ is orthogonal to the $|00\rangle$ state and the same respectively for the other states with their respective orthogonal states. The PBR basis is an orthogonal basis and therefore after the measurement we can eliminate the possibility of the one of the four states from (1.32) with certainty. If we obtain the result corresponding to $|\psi_{\overline{00}}\rangle$ then there is zero possibility of the initial state being $|00\rangle$ and so we have eliminated it.

To check that the PBR basis is complete we can sum up the projectors onto states from (1.33) and if the measurement is complete they will sum to the identity. In vector form we have the following relationships,

$$\begin{aligned}
|00\rangle &= (1,0,0,0)^\dagger, \\
|01\rangle &= (0,1,0,0)^\dagger, \\
|10\rangle &= (0,0,1,0)^\dagger, \\
|11\rangle &= (0,0,0,1)^\dagger.
\end{aligned} \tag{1.34}$$

Using this we can represent the projectors as

$$
\begin{aligned}
|\psi_{\overline{00}}\rangle\langle\psi_{\overline{00}}| =& \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
|\psi_{\overline{0+}}\rangle\langle\psi_{\overline{0+}}| =& \frac{1}{4}\begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{pmatrix}, \\
|\psi_{\overline{+0}}\rangle\langle\psi_{\overline{+0}}| =& \frac{1}{4}\begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}, \\
|\psi_{\overline{++}}\rangle\langle\psi_{\overline{++}}| =& \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned} \tag{1.35}
$$

Summing up all the projectors we get

$$
|\psi_{\overline{00}}\rangle\langle\psi_{\overline{00}}| + |\psi_{\overline{0+}}\rangle\langle\psi_{\overline{0+}}| + |\psi_{\overline{+0}}\rangle\langle\psi_{\overline{+0}}| + |\psi_{\overline{++}}\rangle\langle\psi_{\overline{++}}| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.36}
$$

As the result is the identity we have an orthonormal basis and therefore know that a measurement in the PBR basis will eliminate one of the four states in (1.32) every single time. This is an interesting outcome due to the fact that $|0\rangle$ and $|+\rangle$ are non-orthogonal and so local measurements on the individual qubits would not be able to achieve the same result.

The authors then proceed to consider the state

$$
\begin{aligned}
|\psi_0\rangle &= \cos(\theta/2)|0\rangle + \sin(\theta/2)|1\rangle, \\
|\psi_1\rangle &= \cos(\theta/2)|0\rangle - \sin(\theta/2)|1\rangle.
\end{aligned} \tag{1.37}
$$

For n systems we have 2^n states

$$
|\Psi_{x_1,x_2\ldots x_n}\rangle = |\psi_{x_1}\rangle \otimes |\psi_{x_2}\rangle \otimes ..... \otimes |\psi_{x_n}\rangle, \tag{1.38}
$$

where $x_i \in \{0, 1\}$ for each i. It is then stated that for $0 < \theta < \pi/2$ if you choose n large enough to satisfy

$$
2^{1/n} - 1 \leq \tan(\theta/2), \tag{1.39}
$$

a measurement exists that will eliminate one of the 2^n states with certainty. Checking this in the two qubit case we can see the overlap

$$\langle 0|+\rangle = \frac{1}{\sqrt{2}}, \tag{1.40}$$

and

$$\langle \psi_0|\psi_1\rangle = \cos^2(\theta/2) - \sin^2(\theta/2) = \cos(\theta). \tag{1.41}$$

Therefore $\theta = \pi/4$ is equivalent to the $|0\rangle, |+\rangle$ case. If we insert $\theta = \pi/4$ into equation (1.39) then we obtain

$$2^{1/n} - 1 \leq \tan(\pi/8) = \sqrt{2} - 1. \tag{1.42}$$

If we solve for n, then we obtain

$$\begin{aligned} &2^{1/n} \leq (1 + (\sqrt{2} - 1) = \sqrt{2}, \\ &\frac{1}{n}\ln(2) \leq \frac{1}{2}\ln(2), \\ &\frac{1}{n} \leq \frac{1}{2}, \\ &n \geq 2. \end{aligned} \tag{1.43}$$

The answer is expected as we know for the case with $|0\rangle$ and $|+\rangle$ there exists a two-qubit measurement that eliminates one of the states with certainty.
The elimination measurement is important in proving this no-go theorem as it gives us a result with certainty each time for a measurement with a system of two or more distinct quantum states with an overlap in their probability distributions. If we tried to discriminate between the two qubit states we would not get an unambiguous result every single time.

PBR Experimental Evidence

Experimental tests have been performed to test the PBR theorem [18, 19, 20] and similar no go theorems testing whether ψ-*epistemic* models can explain the indistinguishability of quantum states [21, 22, 23]. The first test by Nigg et al. [18] in 2012 used trapped ions, whereas the more recent experiments [19, 20] have used single photons. The results of the experimental tests support the PBR theorem and the general inadequacy of ψ-*epistemic* models. The experiments don't reproduce the exact measurement system from the PBR paper yet use similar methods to represent the same contradictions.

1.4 Aims of the Thesis

In the previous section I introduced the PBR measurement as the motivation for the work in a large portion of this thesis. We based our work on expanding the results and studying more outcomes than those within the PBR paper.

1.4.1 Notation

Before we state our problems I will introduce the notation used in the following thesis. We define the states we study as

$$\begin{aligned} |\theta\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle, \\ |-\theta\rangle &= \cos\theta|0\rangle - \sin\theta|1\rangle, \end{aligned} \tag{1.44}$$

with respective orthogonal states,

$$\begin{aligned} |\overline{\theta}\rangle &= \sin\theta|0\rangle - \cos\theta|1\rangle, \\ |\overline{-\theta}\rangle &= \sin\theta|0\rangle + \cos\theta|1\rangle. \end{aligned} \tag{1.45}$$

We have chosen to use θ instead of the $\theta/2$ convention for a Bloch vector as it clarified the workings throughout the thesis and we don't refer to the Bloch sphere picture within the measurement work.
If there are two qubits and each one can be in either of the states from (1.44) then the four possible states are

$$|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle, \tag{1.46}$$

where $|\theta,\theta\rangle$ is the product of the two $|\theta\rangle$ states, alternatively written as $|\theta\rangle \otimes |\theta\rangle$.

1.4.2 Target Problems

In the PBR measurement [1] the authors give a measurement that eliminates one two-qubit state with perfect success, and show that perfect success can be achieved for any two possible states for each qubit given you have enough qubits. From this we decided to try and find the measurements to unambiguously eliminate one two-qubit state when the overlap between the two states was less than $1\sqrt{2}$, giving a case when perfect elimination wasn't possible and so a non-zero failure probability is required. This would give us an optimal measurement for unambiguously eliminating one of the four possible preparations $|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle$ for all θ. Following on from this we looked at eliminating more than a single two-qubit state and finding the measurements required in this situation. For example in the two-qubit case eliminating two out of the four possible two-qubit states. Eliminating three out of the four possible states is equivalent to quantum state discrimination and so this becomes similar to the work done on quantum discrimination of sequences of states [24]. We also briefly extended the work to a three-qubit sequence where there are now eight possible initial states as the number of possible preparations is given by 2^N, where N represents the number of qubits.
Another aim was to look at eliminating the highest average number of states and whether

performing a measurement on each qubit in the sequence (local measurement) would eliminate as many states as measuring in an entangled basis on multiple qubits (global measurement).

1.5 Examples Of State Elimination Measurements

In general there has been a lot more research into state discrimination than there has been for state elimination, but as well as the PBR measurement there have been a few other scenarios in which elimination measurements have been researched and used. We shall go through a few examples in the following work.

1.5.1 Mutual Information and Trine States

The trine ensemble has been studied a few times and a good explanation of it is given in [7]. The trine ensemble is composed of the states

$$\begin{aligned} |\psi_1\rangle &= |0\rangle, \\ |\psi_2\rangle &= -\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle, \\ |\psi_3\rangle &= -\frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle. \end{aligned} \tag{1.47}$$

Then an anti-trine measurement can be formed with the measurement operators,

$$\begin{aligned} \pi_1 &= \frac{2}{3}|1\rangle\langle 1|, \\ \pi_2 &= \frac{2}{3}\left(\frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|0\rangle\right)\left(\frac{1}{2}\langle 1| + \frac{\sqrt{3}}{2}\langle 0|\right), \\ \pi_3 &= \frac{2}{3}\left(\frac{1}{2}|1\rangle - \frac{\sqrt{3}}{2}|0\rangle\right)\left(\frac{1}{2}\langle 1| - \frac{\sqrt{3}}{2}\langle 0|\right). \end{aligned} \tag{1.48}$$

This is an elimination measurement where the probability of each outcome is given as

$$p(j|i) = \langle\psi_i|\pi_j|\psi_i\rangle = \frac{1}{2}(1 - \delta_{ij}). \tag{1.49}$$

So if $i = j$ there is zero probability of that outcome occurring, which is the definition of an unambiguous elimination measurement. The $1/2$ factor states that there is an equal probability of the correct result being either of the remaining states once a state has been eliminated. Interestingly this elimination measurement gives the optimal value for mutual information [7]. For a quantum channel the mutual information is given by,

$$H(A:B) = \sum_{ij} p_i \mathrm{Tr}(\rho_i\pi_j)\log\left(\frac{\mathrm{Tr}(\rho_i\pi_j)}{\mathrm{Tr}(\rho\pi_j)}\right), \tag{1.50}$$

where $\rho = \sum_i p_i \rho_i$. p_i are the prior probabilities for the states ρ_i. A labels the initial state and B is the measurement outcome related to the measurement operators π_j. For two pure states it is known the method for maximising the mutual information with states of equal probability is to perform a minimum-error measurement [25]. For the trine states (1.47) with three equally probably states it is the elimination measurement with measurement operators from (1.48), instead of a minimum-error measurement, that maximises the mutual information.

1.5.2 Quantum Digital Signature Implementation

Quantum digital signatures (QDS) can be used for transferable message authentication. The security of digital signatures in [26] is information theoretic and therefore can theoretically withstand any attempt of forgery even against a forger with unlimited computing power including the use of a quantum computer. I will introduce a simple overview of how a QDS protocol occurs.

QDS protocols have a distribution stage followed by a messaging stage. In the distribution stage Alice sends an identical sequence of states to all recipients she will later want to send the message to. There will be a different sequence of states for each possible message. In [26] the possible sent states are $\{|\alpha\rangle, |i\alpha\rangle, |-\alpha\rangle, |-i\alpha\rangle\}$ and the elimination measurement shown in figure 1.1 unambiguously eliminates one, two or three of the initial possibilities. The recipients perform the elimination measurement and store the classical results. In the messaging stage Alice or a forger will send the message along with the classical results relating to the signature. The recipients then compare there classical results to those of the sender. If the recipient has eliminated say $|\alpha\rangle$ as the ith state and the sender puts the result associated with $|\alpha\rangle$ in this position then we know it is a forger and the signature is invalid. This is assuming perfect practical implementation, in reality a threshold for the percentage incorrect would be used. The advantage of using an unambiguous state elimination (USE) measurement as opposed to unambiguous state discrimination is that the success probability for eliminating less than $N-1$ states can be higher than the success probability of USD. The other advantage for the protocol USE measurement in [26] is that if the measurement fails to unambiguously discriminate (eliminate $N-1$ states) then the measurement may still rule out some of the states. Figure 1.1 shows the set-up of the USE used in [26].

Alice initially chooses one state from the four coherent states $\{|\alpha\rangle, |i\alpha\rangle, |-\alpha\rangle, |-i\alpha\rangle\}$ and this state will be labelled $|\beta\rangle$. After interacting with the setup as shown in 1.1 the resulting output state is,

$$|\psi_{out}\rangle = \left|\frac{\beta-\alpha}{2}\right\rangle_1 \otimes \left|\frac{\beta+\alpha}{2}\right\rangle_2 \otimes \left|\frac{\beta-i\alpha}{2}\right\rangle_3 \otimes \left|\frac{\beta+i\alpha}{2}\right\rangle_4, \tag{1.51}$$

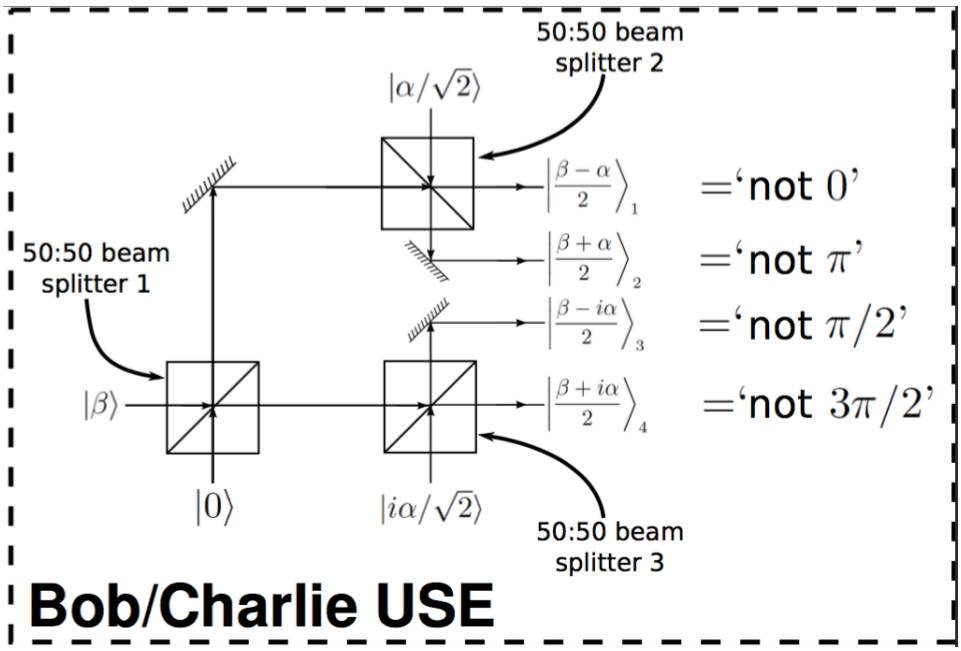


Figure 1.1: This figure is reproduced from [26] and shows the setup for the USE measurement on four coherent states. The incident state $|\beta\rangle$ from Alice is input into a $50/50$ beamsplitter with a vacuum state, then the bottom arm is interfered with a reference beam in the state $|i\alpha/\sqrt{2}\rangle$ and the top beam with a reference beam in the state $|\alpha/\sqrt{2}\rangle$. Clicks in the respective detectors each correspond to an elimination outcome.

where β is the state Alice chose and the α part comes from the reference beams. If we get a click in detector one we know that $|\beta\rangle \neq |\alpha\rangle$. This is because if $|\beta\rangle = |\alpha\rangle$ then we would expect a vacuum state in the first detector. Similarly for the other possible input states $\{|i\alpha\rangle, |-\alpha\rangle, |-i\alpha\rangle\}$. These can also be described by a phase shift from the initial state $|\alpha\rangle$, which is what the $\{0, \pi/2, \pi, 3\pi/2\}$ refer to in figure 1.1. Each detector that clicks refers to the elimination of one of the initial states and it is possible to eliminate one, two or all three of the possibilities available.

1.5.3 Communication Tasks

Perry et al. introduced a game referred to as the exclusion game where quantum resources are infinitely better than their classical counterpart [27]. The game involves Alice, Bob and a referee. Alice is given an n-bit string $\vec{x} \in \{0,1\}^n$ by the referee where each of the 2^n possible n-bit strings are equally likely. Alice can then send a single message to Bob regarding her bit string. The referee then chooses a random subset , $y \subseteq [n]$ of a predetermined size m, of locations in Alice's bit string and gives this to Bob. Bob's task is then to produce an m-bit string that is different to Alice's string in the chosen locations. Bob's aim is then to exclude Alice's m-bit string and if he gives any other m-bit string he

will win the game. For example if $n = 4$ and $m = 2$ and Alice's n-bit string is $\vec{x} = 1010$ and $y = \{2, 4\}$, Bob's winning bit strings are anything except 00 as these are the 2nd and 4th bits of Alice's string. Any of $\{01, 11, 10\}$ will be a winning solution. The measure of success of a protocol is not how often Bob succeeds but in fact how little information Alice has to give to Bob for him to succeed every time. The protocol given in [27] is based on the minimum angle required for elimination from n states derived in [1] given by equation (1.39). Defining the minimum angle to eliminate one state with certainty for m systems as

$$\theta_m = 2 \arctan(2^{1/m} - 1), \tag{1.52}$$

where we have used m instead of n as in the protocol the aim is to eliminate from the subset of states y of size of m. The authors of [27] lay out the protocol and information cost in the supplementary information. The protocol goes as follows:

1. Alice receives the bit string $\vec{x}$ from the referee.

2. Alice prepares the state:

$$|\Psi_{\vec{x}}(\theta_m)\rangle = \bigotimes_{i=1}^{n} |\psi_{x_i}(\theta_m)\rangle, \tag{1.53}$$

   where $|\psi_{x_i}(\theta_m)\rangle$ is either $|\psi_0\rangle$ or $|\psi_1\rangle$ from (1.37) depending on the respective bit from $\vec{x}$.

3. Alice sends $|\Psi_{\vec{x}}(\theta_m)\rangle$ to Bob.

4. On the m systems specified by y Bob performs the measurement to eliminate one of the possible m-qubit states then gives this state to the referee.

We know this is a winning strategy as there is a 100% success rate for the measurement. The authors of then proceed to calculate the entropy of the message (state) sent by Alice. For large m

$$S(|\Psi_{\vec{x}}(\theta_m)\rangle) < \frac{n}{m^2}(2 \ln 2)^2 \left[\frac{1}{\ln 2} + \log_2 \left(\frac{m^2}{(\ln 2)^2} \right) \right], \tag{1.54}$$

so provided $m > n^{\frac{1}{2}} + \beta$, where $\beta > 0$, then the entropy of the message sent by Alice tends to zero in the limit of large n. The authors also show that for a classical strategy Alice is required to send nearly n bits of information about $\vec{x}$ to Bob. This is how an elimination measurement is used to give infinite quantum-classical separation.

1.6 Conclusion

In this introduction we have shown the formalism of quantum measurements, providing information on how quantum mechanics can be used to analyse systems. As well as this we have introduced the concept of state elimination measurements, with some examples of cases when elimination measurements have been used to either optimise certain requirements or for communication theorems. The PBR theorem was also shown and how this utilises an elimination measurement to present it's no-go theorem for epistemic models of quantum mechanics. The elimination measurement used in the PBR theorem became a starting point for our work in the following chapters and we attempt to find optimal measurements for more generalised versions of the PBR measurement. Next we present the numerical work involved in attempting to find the optimal success probabilities for elimination measurements.

Chapter 2

Semi-Definite Programming

2.1 Computational Approach

We started with an analytic approach of finding the highest success probabilities and optimal measurements that gave those probabilities. One of the problems is that we found measurements that we thought would be good but proving they were optimal was not so easy. Therefore we started looking for an algorithm that could produce optimal results so we could test our analytic approaches.

On the computational side we started with the idea of a brute-force search of the space in which the measurement operators lie, with constraints to limit the number of free variables. As explained below the space had eight complex variables so required a search over sixteen parameters and upon first inspection it took too long to find an accurate solution to the problem in a feasible time period.

Semi-definite programming (SDP)[28] is a form of convex optimisation that is suited to a variety of problems. It is similar to the more well known linear programming that is studied in mathematics before university level, yet SDP has only been used more extensively in the last twenty years. Any linear program can be written in the form of a semi-definite program due to the fact the positive quadrant restriction for linear programming lies within the semi-definite realm. This method is well suited to the formalism of quantum mechanics, which is also based largely on semi-definite components. The disadvantage of this approach is that often you only learn a bound on the success probability and not the exact result.

2.2 Brute Force

For the first attempt at finding an exact solution we looked at whether it was possible to do an exhaustive search of the Hilbert space to find the optimal measurements.
Initially we were looking at finding the highest probability to eliminate one of the four states in the two-qubit case. For a brute-force approach we started from the fact that for an unambiguous measurement, the measurement must project onto a state that is orthogonal to the input state we desire to eliminate. The condition for eliminating a pure state $|\psi\rangle$ unambiguously is

$$p(\text{not}|\psi\rangle||\psi\rangle) = \langle\psi|\Pi_{\text{not }\psi}|\psi\rangle = 0. \tag{2.1}$$

That is the probability of the outcome not$|\psi\rangle$ given the input state was $|\psi\rangle$ must be zero. For the state $|\theta,\theta\rangle$ the measurement basis must be composed of some linear combination of $|\overline{\theta},\theta$, $|\theta,\overline{\theta}\rangle$ and $|\overline{\theta},\overline{\theta}\rangle$.
So each possible state has another general state that is orthogonal to it and these are

$$\begin{aligned}
|\psi_{\overline{\theta,\theta}}\rangle &= \alpha_1|\overline{\theta},\theta\rangle + \alpha_2|\theta,\overline{\theta}\rangle + \alpha_3|\overline{\theta,\theta}\rangle,\\
|\psi_{\overline{\theta-,\theta}}\rangle &= \beta_1|\overline{\theta},-\theta\rangle + \beta_2|\theta,\overline{-\theta}\rangle + \beta_3|\overline{\theta,-\theta}\rangle,\\
|\psi_{\overline{-\theta,\theta}}\rangle &= \gamma_1|\overline{-\theta},\theta\rangle + \gamma_2|-\theta,\overline{\theta}\rangle + \gamma_3|\overline{-\theta,\theta}\rangle,\\
|\psi_{\overline{-\theta,-\theta}}\rangle &= \delta_1|-\theta,\overline{-\theta}\rangle + \delta_2|-\theta,\overline{-\theta}\rangle + \delta_3|\overline{-\theta,-\theta}\rangle,
\end{aligned} \tag{2.2}$$

where $|\psi_{\overline{\theta,\theta}}\rangle$ represents the state orthogonal to $|\theta,\theta\rangle$ and similarly for the other states. The normalisation conditions are

$$\begin{aligned}
|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 &= 1,\\
|\beta_1|^2 + |\beta_2|^2 + |\beta_3|^2 &= 1,\\
|\gamma_1|^2 + |\gamma_2|^2 + |\gamma_3|^2 &= 1,\\
|\delta_1|^2 + |\delta_2|^2 + |\delta_3|^2 &= 1.
\end{aligned} \tag{2.3}$$

These conditions can reduce the number of variables by setting $\alpha_3 = \sqrt{1-|\alpha_1|^2-|\alpha_2|^2}$ and similarly for β_3, γ_3 and δ_3. These can be chosen as real without loss of generality. So now we have eight variables but as each of these is complex this leads to having to calculate the success probabilities whilst varying sixteen different parameters.

From this we go through the process of finding the success probability using the method described in chapter 1.2.5 using the the sum of the projectors onto the states given in (2.2), shown as

$$|\psi_{\overline{\theta,\theta}}\rangle\langle\psi_{\overline{\theta,\theta}}| + |\psi_{\overline{\theta-,\theta}}\rangle\langle\psi_{\overline{\theta-,\theta}}| + |\psi_{\overline{-,\theta\theta}}\rangle\langle\psi_{\overline{-\theta,\theta}}| + |\psi_{\overline{-\theta,-\theta}}\rangle\langle\psi_{\overline{-\theta,-\theta}}|. \tag{2.4}$$

This is then scaled using the weightings $\alpha_i, \beta_i, \gamma_i$ and δ_i to make sure the measurement operators sum to the identity. This is followed by calculating the success probability for

eliminating one of four states. If the probability is the highest so far you store that value and the measurement operators and repeat the process.

One method to search the parameter space was we run through each variable eleven times increasing by 0.1 each time from 0 to 1, this would take 11^{16} runs and the precision is very poor. It did not seem immediately feasible to run the program to the desired level of precision of at least three significant figures. One idea would be to do a search with low precision to begin with then study around that area with higher precision, yet it is possible we will just fall into a local minimum. At this point we decided to look into using semi-definite programming (SDP). This seemed a promising tool so we started investigating it's potential for our elimination measurements. First of all to have a look at convex optimisation we will briefly look at linear programming as it is the more common version of convex optimisation studied.

2.3 Linear Programming

Linear programming is analogous to SDP and is a useful introduction to the methods involved as well as being something that occurs more frequently in mathematical education, so is therefore a good starting point. The process of convex optimisation requires the data to be confined to a convex cone. The definition of a convex cone is

Definition 1 *Let E be a real vector space. If P is a subset of E, a vector of the form $\boldsymbol{y} = \sum_{\boldsymbol{x}\in P} \lambda_{\boldsymbol{x}}\boldsymbol{x}$ where $\lambda_{\boldsymbol{x}} \geq 0$ for all $\boldsymbol{x} \in P$ is called a 'positive linear combination' of the elements of P. A subset P of E is a convex cone is P contains all positive linear combinations of any pair of its vectors.*

In two dimension this simplifies, so that if we have $\lambda_1\mathbf{x} + \lambda_2\mathbf{y}$ then P is a convex cone for any $\lambda_1, \lambda_2 \geq 0$ and any **x**,**y** in P.

Linear programming is convex optimisation in which the convex cone is the non-negative orthant ($\mathbf{x}_1 \geq 0, \mathbf{x}_2 \geq 0 ... \mathbf{x}_n \geq 0$), for example in 2D this is just the positive quadrant. The cone just represents the space that the variables must belong to. Simply linear programming is an optimisation problem when the objective function, variables and constraints involved are linear.

A linear program in the canonical form can be expressed as

$$\begin{aligned} \text{Maximise} \quad & \mathbf{a}^T\mathbf{x}, \\ \text{subject to} \quad & C\mathbf{x} \leq \mathbf{b}, \\ \text{and} \quad & \mathbf{x} \geq 0. \end{aligned} \tag{2.5}$$

Here **a**, and **x** are fixed and **x** is the variable we alter to maximise $\mathbf{a}^T\mathbf{x}$. The conditions are written out in vector form and so the inequality $C\mathbf{x} \leq \mathbf{b}$ is done by checking elementwise that $\mathbf{b} - C\mathbf{x} \geq 0$. Each element represents a condition and this can be seen in the following example.

2.3.1 A Simple Example

You have decided to open a fruit stand. You have $100m^2$ of land available to plant either gooseberries (G) or strawberries (S). From $1m^2$ you can grow enough gooseberries to sell for £2 and enough strawberries to sell for £4. Due to water restrictions you are only allowed 150 gallons of water in total for the $100m^2$, gooseberries and strawberries require one and three gallons per m^2 respectively.

$$\begin{aligned} \text{Maximise} \quad & 2G + 4S, \\ \text{subject to} \quad & G + S \leq 100, \\ & G + 3S \leq 150, \\ \text{and} \quad & G, S \geq 0. \end{aligned} \tag{2.6}$$

The first line gives the optimisation problem which in this case is to gain the largest income. The first and second conditions are from the size and water restrictions respectively. Then finally there is the non-negativity requirement. To relate this to the standard formula from (2.5) we have

$$\mathbf{a} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} G \\ S \end{pmatrix}, \ C = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 100 \\ 150 \end{pmatrix}. \tag{2.7}$$

All linear programs and thus semidefinite programs can fit into the formulation above. There are many ways to solve simple linear programs [29] for example , graphical methods, simplex method and the least cost method to name a few.

2.4 Introduction To Semidefinite Programming

Definition 2 *A semi-definite program is a triple $(\phi(X), A, X)$, with the aims to*

$$\begin{aligned} \textit{Maximise} : \quad & \langle A, X \rangle \\ \textit{subject to} : \quad & \phi(X) = B \\ \textit{and} \quad & X \succeq 0. \end{aligned} \tag{2.8}$$

where $\langle A, X \rangle$ represents the inner product of A and X.

In this program A is fixed along with the components of the conditions $\phi(X)$ and B. X is the variable in the program we are optimising over and A, B and X are all matrices. This program is very similar to that of linear programming from (2.5), yet in a semi-definite program X lies in the cone of positive semi-definite matrices whereas in linear programming X just lies in the non-negative orthant.

Definition 3 *A hermitian matrix X is positive semi-definite (PSD) iff all the eigenvalues of X are non-negative.*

Quantum measurements fit perfectly into the framework of SDP. For example, a POVM is a set of positive semidefinite operators and they can act in a linear manner on another positive semidefinite entity such as a density operator. They are subject to linear conditions such as the requirement that the sum of all measurement operators should be the identity operator. For this reason semi-definite programs have been used extensively in a variety of ways related to quantum mechanics including measurements [30], quantum error correction [31] and quantum cryptography [32]. The method has been used for state elimination in the PBR paper [1] to give a numerical result to complement the analytical one and also by Bandyopadhyay et al.[9] in the study of conclusive exclusion of quantum states. For this reason we chose to use SDP as the numerical tool to obtain the optimal success probabilities for elimination measurements and to find the bounds on the minimum separation required of the quantum states for certain elimination (success probability of one). Some useful resources for information on SDP are the lecture notes by John Watrous [33] and also by Jamie Sikora [34].

In convex optimisation it is possible to formulate a dual problem from which one can produce a bound on the optimal value from the primal problem. The primal and dual problems are mathematically related to each other as will be shown in the following work. Often the dual problem can be easier to solve and so the bounds on the primal problem can be obtained without having to solve a potentially difficult primal problem. This is not always the case though and it is definitely worth attempting to solve the primal problem directly before formulating a dual problem. In most literature the primal problem is introduced alongside a dual problem and sometimes the dual problem is not even required. For quantum measurements most primal problems and dual problems can be formulated as

Primal Problem	***Dual Problem***
supremum (sup) $Tr[AX]$.	infimum (inf) $Tr[BY]$.
Subject to $\phi(X) = B$,	Subject to $\phi^*(Y) \geq A$,
$X \geq 0$.	$Y \in$ Herm,

(2.9)

where Herm is the set of hermitian matrices and $\phi^*(Y)$ is the map which is the dual of $\phi(X)$. The dual is given by $Tr[Y\phi(X)] = Tr[X\phi^*(Y)]$. Infimum and supremum refer to the aim of finding the infimum or supremum of the problem. As convex optimisation does not guarantee to find the optimal result then we may not be able to find the maximum or minimum so supremum and infimum are used as the best results of those accessible to convex optimisation. If we prove the results obtained are optimal by methods explained later then we can user the terms minimum and maximum.
The primal and dual optimal values are respectively given by:

$$\alpha^* = \sup_{X\in\mathcal{A}} Tr[AX] \qquad \text{and} \qquad \beta^* = \inf_{Y\in\mathcal{B}} Tr[BY]. \tag{2.10}$$

The dual problem is related to the primal problem and can be used to obtain information about the solution to the primal problem. As the dual problem is also a linear program you can of course also find the dual of the dual problem, yet all this will do is take you back to the primal problem again.

2.5 Minimum-Error Quantum State Discrimination SDP

Quantum state discrimination (QSD) is a well studied field and there are many cases where we know the optimal measurements and related success probabilities. For example there is the Helstrom bound [35] that gives the minimum probability of making an error as

$$P_{err} = \frac{1}{2}\left(1 - \sqrt{1 - 4p_0p_1|\langle\psi_0|\psi_1\rangle|^2}\right), \tag{2.11}$$

where p_0 and p_1 are the probabilities of the states $|\psi_0\rangle$ and $|\psi_1\rangle$ respectively.
I will now go through formulating a semi-definite program for minimum-error discrimination of a set of states.
Table 2.1 describes the components involved in minimum-error QSD and how they relate to the SDP notation used in (2.9), I came across this table for general SDP formulation in a talk from Jamie Sikora [36] and it seemed a useful way to formulate the SDP. The first and second column are generic to most semi-definite programs. We don't always have states and POVMS but in general the problem can be split into data, variables, constraints and the aim. The third column however is specific for minimum-error QSD. In table 2.1 p_i are the a priori probabilities of the state represented by the density matrix ρ_i being sent, with M_i being the measurement operators.
To formulate the dual problem we use the definition of the adjoint $Tr[Y\phi(X)] = Tr[X\phi^*(Y)]$ along with the knowledge $\phi(X) = \sum_i M_i$ and $X = M_i$ to obtain

$$Tr[Y\textstyle\sum_i M_i] = Tr[M_i\phi^*(Y)]$$

Property	SDP	QSD
Data (States)	A	$\{p_1\rho_1, ..., p_n\rho_n\}$
Variables (POVMS)	X	$\{M_1, ..., M_n\}$
Constraints	$\phi(X) = B$	$\sum_{i=1}^{n} M_i = I$
Aim	$sup\ Tr[AX]$	sup $\sum_{i=1}^{n} p_i Tr[\rho_i M_i]$

Table 2.1: Table showing the relationship between SDP notation from (2.9) and the program for the minimum-error discrimination measurement of quantum states.

Dual notation	QSD
Y	Y
B	I
$\phi^*(Y)$	Y
A	$p_i\rho_i$

Table 2.2: Table showing the equivalence of the notation in our problem with that from equation (2.9).

$$Tr\left[Y\sum_i M_i\right] = Tr\left[\begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_n \end{pmatrix}\begin{pmatrix} Y & & \\ & \ddots & \\ & & Y \end{pmatrix}\right], \tag{2.12}$$

where the empty spaces denote zeroes in those positions, for example if there are three 2X2 M_i matrices then the matrix would be

$$\begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_3 \end{pmatrix} = \begin{pmatrix} M_{1(11)} & M_{1(12)} & 0 & 0 & 0 & 0 \\ M_{1(21)} & M_{1(22)} & 0 & 0 & 0 & 0 \\ 0 & 0 & M_{2(11)} & M_{2(12)} & 0 & 0 \\ 0 & 0 & M_{2(21)} & M_{2(22)} & 0 & 0 \\ 0 & 0 & 0 & 0 & M_{3(11)} & M_{3(12)} \\ 0 & 0 & 0 & 0 & M_{3(21)} & M_{3(22)} \end{pmatrix}. \tag{2.13}$$

We shall use this notation throughout to simplify the equations. Looking back to equation (2.12), by defining a generic matrix *Y* for the variable in the dual we can then set $\phi^*(Y)$ as a block diagonal matrix with *Y* matrices along the diagonal. By doing this we have satisfied the adjoint condition and can formulate the dual program by relating the SDP dual problem notation from (2.9) to the QSD case as described in table 2.2. From this table we can formulate the primal and dual problems.

Primal Problem

supremum $\sum_i Tr[p_i \rho_i M_i]$.
Subject to $\sum_i^n M_i = I$,
$M_i \geq 0 \ \ \forall i$.

Dual Problem

infimum $Tr[Y]$.
Subject to $Y \geq p_i \rho_i \ \forall i$,
$Y \in$ Herm.

2.6 Duality Bounds

With the dual problem we can obtain an upper bound for the primal problem. This is known as weak duality and can be stated as

Theorem 1 (Weak Duality) $\alpha \leq \beta$. *Any feasible solution to the dual problem is a lower bound for the primal problem and vice versa.*

α and β are optimal solutions to the primal problem and dual problem respectively. The theorem can be proven easily as shown below,

$$\langle A, X \rangle - \langle B, Y \rangle = \langle A, X \rangle - \langle \phi(X), Y \rangle = \langle A, X \rangle - \langle \phi^*(Y), X \rangle = \langle A - \phi^*(Y) \rangle \leq 0, \tag{2.14}$$

as $\phi^*(Y) \geq A$. This leads to

$$\langle A, X \rangle - \langle B, Y \rangle \leq 0 \qquad \text{and} \qquad \alpha \leq \beta. \tag{2.15}$$

The above holds for a maximisation problem such as the quantum state discrimination problem described above. For a minimisation problem the inequality in the theorem just switches from a less than or equal to greater than or equal leading to $\alpha \geq \beta$. The difference between the optimal solutions of the primal and dual problems is called the duality gap. If the duality gap equals 0 then we have strong duality, otherwise we have weak duality. The condition for strong duality in convex optimisation programs was discovered by Slater [37] and is hence known as Slater's condition.

Theorem 2 (Slater's condition) *Given a semidefinite program in the standard form we have the primal components $\phi(X), X, A$ and B. Suppose the feasible set of primal solutions is P and the feasible set of dual solutions is D. A set of solutions is feasible if they satisfy the conditions the primal and dual problems are subject to. Strong duality holds if either*

-$D \neq \emptyset$ and there exists a strictly feasible $X \in P, \ i.e. \ X > 0, \ \phi(X) = B$

-$P \neq \emptyset$ and there exists a strictly feasible $Y \in D, \ i.e. \ \phi^*(Y) - A \geq 0$.

In most interesting cases the feasible set is non-empty and in that case we know $D \neq \emptyset$ and $P \neq \emptyset$, therefore we are more looking to find strictly feasible parameters to test for strong

duality. In the case of semidefinite programming the requirement to be strictly feasible is that must have $X > 0$ and so require positive definite matrices instead. For a positive semi-definite matrix we just required the matrix to be hermitian and have non-negative eigenvalues, whereas for it to be positive definite it must now have positive eigenvalues. All measurement operators are either positive semidefinite or positive definite as for any measured state we require the outcome to have a non-negative probability.
As well as Slater's condition we have other methods to study the duality of the programs.

Theorem 3 (Complementary Slackness) *Suppose the optimal solution to the primal problem is $\tilde{X}$ and the dual is $\tilde{Y}$ and that $\alpha = \beta$. Then*

$$\phi^*(\tilde{Y})\tilde{X} = A\tilde{X} \qquad \textit{and} \qquad \phi^*(\tilde{X})\tilde{Y} = B\tilde{Y} \tag{2.16}$$

This can be shown as for an optimal solution we have

$$\langle A, X\rangle = \langle B, Y\rangle = \langle \phi(x), Y\rangle = \langle \phi^*(Y), X\rangle, \tag{2.17}$$

so

$$\langle \phi^*(Y) - A, X\rangle = 0.$$

The inner product of the two semidefinite matrices $\phi^*(Y) - A$ and X is zero if and only if their product is zero. Therefore we get,

$$(\phi^*(Y) - A)X = 0 \tag{2.18}$$

and from this we can obtain the first equation from (2.16) and the second just comes from the fact $\phi(x) = B$. With this it can be shown there is a unique operator Y that is an optimal dual solution [33]. With this theorem you can check whether a primal result is optimal. This can be done by giving no slack ($\phi^*(Y) = A$) on the dual constraint and from this you can calculate the optimal dual solution and then if this is feasible it will be the optimal primal solution. Note this only works for strong duality as you require $\langle A, X\rangle = \langle B, Y\rangle$.

2.7 Quantum State Elimination

In quantum state elimination, measurement outcome i is associated with ruling out state ρ_i. Our aim is, broadly speaking, to minimise the probabilities to obtain outcome i if the state is ρ_i. An ideal case is $Tr[\rho_i M_i] = 0 \; \forall \; i$ with the measurement operators orthogonal to the state or set of states we wish to eliminate and therefore giving us an unambiguous result. Otherwise we may wish to maximise the probability to guess correctly which state was eliminated, this would then just be a minimum error elimination measurement.

Elimination semi-definite programs have been investigated before [9] with the main aim of eliminating a single state with certainty. We are also looking more into multiple state elimination where the aim is to eliminate more than one state with a single measurement. We also studied inconclusive measurements, when the probability of success $P_s < 1$. First of all we will look at minimum-error quantum state elimination (QSE). We can see

Property	SDP	QSE
States	A	$\{p_1\rho_1, ..., p_n\rho_n\}$
POVM	X	$\{M_1, ..., M_n\}$
Constraints	$\phi(X) = B$	$\sum_{i=1}^{n} M_i = I$
Aim	inf α	inf $\sum_{i=1}^{n} p_i Tr[\rho_i M_i]$

Table 2.3: Table showing the relationship between the standard form for SDP and the form for the minimum-error QSE program.

from table 2.3 that the form of the program is very similar to that of the quantum state discrimination measurements shown in table 2.1. In fact the only change is now we have a minimisation problem instead of a maximisation. Therefore calculating the adjoint $\phi^*(Y)$ will be the same just with the infimum and supremum swapped as well as the inequality on the constraint in the dual problem.

Primal Problem

inf $\sum_i Tr[p_i\rho_i M_i]$.
Subject to $\sum_i^n M_i = I$,
$M_i \geq 0 \ \ \forall i$.

Dual Problem

sup $Tr[Y]$.
Subject to $Y \leq p_i\rho_i \ \forall i$,
$Y \in$ Herm.

For a minimisation problem the optimal solution to the dual problem now gives a lower bound on the optimal solution of the primal problem.

2.7.1 Strong Duality Of The Min-Error QSE Semi-Definite Program

For Slater's condition to hold we require $M_i > 0 \ \forall i$ and $\sum_i^n M_i = I$ to satisfy the first condition of theorem (2) and also $Y \leq p_i\rho_i \ \forall i$ so that the result is feasible.
Taking $M_i = \frac{1}{n}I$ and $Y = -I$ then M_i is positive definite and so strictly feasible and also satisfies the constraint $\sum_i^n M_i = I$. Furthermore $p_i\rho_i$ must be positive as p_i are probabilities and ρ_i are at least positive semidefinite. As Y is negative then $Y \leq p_i\rho_i \ \forall i$. So we have strong duality and therefore $\alpha = \beta$ and solving our dual problem will give us an optimal solution for the primal.

2.8 Eliminating More Than One State

The problems discussed previously involved trying to eliminate one of the initial states. What if we wanted to eliminate two or more states? This was briefly covered by Bandyopadhyay et al. [9], who showed how eliminating multiple states can be expressed in the same form as single state elimination with different initial states. If there are n initial states and we aim to eliminate m of them there are now $\binom{n}{m}$ measurement operators. When we reach the point of $m = n - 1$ we have state discrimination, (instead of n measurement operators we now have $\binom{n}{m}$). For example if there are four initial possibilities A,B,C and D that represent states ρ_A and so on. There are $\binom{4}{2} = 6$ ways of eliminating two of them. These are as follows,

$$AB, AC, AD, BC, BD \text{ and } CD$$

where AB represents eliminating A and B. The aim for a minimum-error measurement is then to minimise,

$$\sum_{i,j=A,B,C,D} Tr[(p_i\rho_i + p_j\rho_j)M_{ij}] \qquad \forall\, i \neq j. \tag{2.19}$$

If we redefine the state $p_i\rho_i + p_j\rho_j$ as $p_k\tilde{\rho}_k$ then we can use the same notation as for single state elimination except we now have $k = \binom{n}{m}$ states instead of m. Therefore the primal and dual programs are formed in the same way except now with the modified states $\tilde{\rho}_k$.

Primal Problem

inf $\sum_i Tr[p_i\tilde{\rho}_i M_i]$.
Subject to $\sum_i M_i = I$,
$M_i \geq 0 \ \ \forall i$.

Dual Problem

sup $Tr[Y]$.
Subject to $Y \leq p_i\tilde{\rho}_i \ \ \forall i$,
$Y \in$ Herm.

This change in states has also caused the number of constraints in the dual problem to become $\binom{n}{m}$ instead of n as before.

2.9 Unambiguous State Elimination (USE)

2.9.1 The USE Semidefinite Program

Minimum-error elimination gives the highest probability of the obtained result being correct. In unambiguous measurements you produce a result that you are certain is correct. This comes at the cost of there being a probability of having an inconclusive or "failure" outcome.

Instead of minimising $Tr[\rho_i M_i]$ we now require $Tr[\rho_i M_i] = 0$ but have $n + 1$ results with a measurement operator M_f that corresponds to an inconclusive result. The aim is

Property	QSE	SDP
Data	$\{p_1\rho_1, ..., p_n\rho_n\}$	A
Variables	$\{M_1, ..., M_n\}$	X
Constraints	$\sum_{i=1}^{n} M_i \leq I$	$\phi(X) = B$
	$Tr[\rho_i M_i] = 0$	
Aim	inf $\sum_{i=1}^{n} p_i Tr[\rho_i M_f]$	inf α

Table 2.4: Table showing the relationship between general SDP notation from (2.9) and the USE measurement of quantum states.

to minimise the probability of obtaining an inconclusive result. Below is the method used by Bandyopadhyay et al. [9]. Table 2.4 shows the components of a USE measurement and how this fits into the SDP format, where $p_i\rho_i$ represent the initial states and there a priori probabilities. The measurement operators M_i are the variables in the problem and correspond to an outcome of unambiguous elimination of the state ρ_i. We also have a measurement operator M_f that corresponds to an inconclusive result and is there to complete the measurement so is defined as,

$$M_f = I - \sum_i M_i. \tag{2.20}$$

This table above is very similar to that of the minimum-error QSE. The only differences are the extra constraints $Tr[\rho_i M_i] = 0$ and there is an inequality in the constraints so now the primal problem isn't in the form of a standard semi-definite program. This is something not addressed by Bandyopadhyay et al. [9]. The solver I use [38] is capable of dealing with constraints in the form of inequalities yet I shall show how we can adapt the constraint to the standard SDP form and then how this won't affect the dual problem produced.

To transform the inequality $\sum_{i=1}^{n} M_i \leq I$ to an equality in line with the standard form we can introduce a slack variable. In this case our slack variable is the inconclusive measurement operator M_f leading to $\sum_{i=1}^{n} M_i + M_f = I$, where $M_f \succeq 0$.

If we take the conditions

$$\begin{aligned} &\sum_{i=1}^{n} M_i \leq I, \;\; I \in \text{Herm}(\mathcal{Y}_1), \\ &Tr[\rho_i M_i] = 0, \;\; 0 \in \text{Herm}(\mathcal{Y}_2), \\ &M_f \in Pos(\mathcal{Y}_1), \end{aligned} \tag{2.21}$$

then the map ϕ becomes

$$\phi \begin{pmatrix} M_f & Y \\ Y^* & X \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} M_i + M_f & 0 \\ 0 & Tr[\rho_i M_i] \end{pmatrix} \tag{2.22}$$

The mapping changes from $\phi : \text{Herm}(\mathcal{X}) \rightarrowtail \text{Herm}(\mathcal{Y})$ to $\phi : \text{Herm}(\mathcal{X} \oplus \mathcal{Y}_1) \rightarrowtail \text{Herm}(\mathcal{Y}_1 \oplus \mathcal{Y}_2)$. Now we can write the semi-definite program in standard form as,

$$\inf \quad \sum_i Tr[(p_i\rho_i \oplus 0)M_i], \tag{2.23}$$

$$\text{Subject to} \quad \phi(X) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.24}$$

$$M_i \in \text{Pos}(\mathcal{X} \oplus \mathcal{Y}_1), \tag{2.25}$$

therefore we can proceed in the normal form for a semi-definite program and not worry about the fact that there is an inequality in our primal constraint. A is a $d \otimes n$ block-diagonal matrix where d is the dimension of the states and n is the number of states, given as

$$A = \begin{pmatrix} \sum_{i=1}^n p_i\rho_i & & \\ & \ddots & \\ & & \sum_{i=1}^n p_i\rho_i \end{pmatrix}. \tag{2.26}$$

X is also a $d \otimes n$ block-diagonal matrix given as

$$X = \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_n \end{pmatrix}. \tag{2.27}$$

To combine the two constraints $\phi(X)$ is defined as

$$\phi(X) = \begin{pmatrix} \sum_i^n M_i & & & \\ & Tr[\rho_1 M_1] & & \\ & & \ddots & \\ & & & Tr[\rho_n M_n] \end{pmatrix}, \tag{2.28}$$

and B as

$$B = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.29}$$

The identity is a $d \otimes d$ matrix and appears due to the requirement that the measurement operators are constrained by the identity and zeroes correspond to the need for an unambiguous elimination measurement outcome to have zero probability of success for the respective measurement operators and states.

To find the adjoint $\phi^*(Y)$ we have

$$Tr[Y\phi(X)] = Tr[X\phi^*(Y)], \tag{2.30}$$

$$Tr\left[Y\begin{pmatrix}\sum_i^n M_i & & & \\ & Tr[\rho_1 M_1] & & \\ & & \ddots & \\ & & & Tr[\rho_n M_n]\end{pmatrix}\right] = Tr\left[\begin{pmatrix}M_1 & & & \\ & \ddots & & \\ & & & M_n\end{pmatrix}\phi^*(Y)\right]. \tag{2.31}$$

If we then define Y as

$$Y = \begin{pmatrix}N & & & \\ & a_1 & & \\ & & \ddots & \\ & & & a_n\end{pmatrix}, \tag{2.32}$$

where a_i is a real variable. To satisfy equation (2.30) we require

$$\phi^*(Y) = \begin{pmatrix}N + a_1\rho_1 & & & \\ & N + a_2\rho_2 & & \\ & & \ddots & \\ & & & N + a_n\rho_n\end{pmatrix}, \tag{2.33}$$

and this gives us the connected pair of programs

Primal Problem	***Dual Problem***
sup $\sum_i Tr[p_i\rho_i M_i]$.	inf $Tr[N]$.
Subject to $\sum_i M_i \leq I$,	Subject to $N + a_i\rho_i \geq \sum_{j=1}^n \rho_j$,
$Tr[\rho_i M_i] = 0$,	$a_i \in \mathbf{R}\ \ \forall i$,
$M_i \geq 0\ \ \forall i$.	$N \in$ Herm.

(2.34)

2.9.2 Duality Of The Unambiguous Programs

For strong duality we require a strictly feasible solution and thus $M_i > 0$ and $\phi(X) = B$. Let us first check with the same approach as we used for the minimum-error case.
Take $M_i = \frac{1}{n}I$ and $Y = -I$. We still have M_i as positive definite except now $\sum_i^n M_i \neq I$ in every case so we might require an operator M_f to complete the measurement. As in our case the sum of the measurement operators we are searching over is not required to be identity due to the existence of the failure operator. This leads to us having weak duality and the solution to the dual problem is an upper bound on the success probability. We can also check the complementary slackness given in theorem 3 .
If the solutions are optimal and strong duality holds they would satisfy the condition of complementary slackness given by,

$$\Phi^*(Y)X = AX \qquad \text{or} \qquad \Phi(X)Y = BY \tag{2.35}$$

With our program this becomes,

$$\begin{pmatrix} N + a_1\tilde{\rho_1} & & \\ & \ddots & \\ & & N + a_k\tilde{\rho_k} \end{pmatrix} \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_k \end{pmatrix} = \begin{pmatrix} \sum \tilde{\rho_j} & & \\ & \ddots & \\ & & \sum \tilde{\rho_j} \end{pmatrix} \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_k \end{pmatrix}. \tag{2.36}$$

This leads to requiring

$$(N + a_i\tilde{\rho_i})M_i = \sum_{j=1}^{k} \tilde{\rho_j} M_i. \tag{2.37}$$

If we take a sum over i over both sides,

$$N \sum_{i=1}^{k} M_i + \sum_{i=1}^{k} a_i\tilde{\rho_i} M_i = \sum_{j=1}^{k} \tilde{\rho_j} \sum_{i=1}^{k} M_i, \tag{2.38}$$

then we see that,

$$N \neq \sum_{j=1}^{k} \tilde{\rho_j} \sum_{i=1}^{k} M_i. \tag{2.39}$$

In this case complementary slackness does not hold for an optimal solution, confirming $\alpha \neq \beta$. We can also look at $\Phi(X)Y = BY$, which gives

$$\begin{pmatrix} \sum_{i=1}^{k} M_i & & & \\ & Tr[\tilde{\rho_1}M_1] & & \\ & & \ddots & \\ & & & Tr[\tilde{\rho_k}M_k] \end{pmatrix} \begin{pmatrix} N & & & \\ & a_1 & & \\ & & \ddots & \\ & & & a_k \end{pmatrix} = \begin{pmatrix} I & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} \begin{pmatrix} N & & & \\ & a_1 & & \\ & & \ddots & \\ & & & a_k \end{pmatrix}, \tag{2.40}$$

leading to,

$$N \sum_{i=1}^{k} M_i = N \qquad \text{and} \qquad a_i \text{Tr}[\tilde{\rho_i} M_i] = 0. \tag{2.41}$$

The right hand equation above is correct due to the fact that $\rho_i M_i = 0$, yet for the left hand equation $\sum_{i=1}^{k} M_i$ does not have to equal the identity in an unambiguous case. In fact when $\sum_{i=1}^{k} M_i = I$ the solution will be the same as in the minimum error case and so we know at this point we have strong duality.

2.9.3 Duality Gap For Unambiguous State Discrimination

If we look at unambiguous state discrimination between the states $|+\theta\rangle$ and $|-\theta\rangle$, we can compare the results obtained from SDP with the known optimal success probability, for distinguishing between these two states. The known optimal success probability is given by the IDP limit that was derived and named after Ivanovic, Dieks and Peres [11][12][13]. For unambiguous discrimination between $|\pm\theta\rangle$ the IDP limit gives the maximum success probability as $1 - \cos(2\theta)$. By finding the difference between the value the SDP gives us and the known optimal probability, this difference will be the duality gap for unambiguous state discrimination. From this we can further verify that strong duality does not hold for unambiguous measurements, where a failure operator is required to complete the measurement. For unambiguous state discrimination the primal problem can be written as

$$\underline{\textbf{\textit{Primal Problem}}}$$

$$\begin{aligned} &\sup \sum_{i=1}^{n} Tr[p_i \rho_i M_i]. \\ \text{Subject to} &\sum_{i=1}^{n} M_i \leq I, \\ &Tr[\rho_i M_j] = 0 \,\forall\, i \neq j, \\ &M_i \geq 0 \;\; \forall i. \end{aligned} \tag{2.42}$$

As for the example with unambiguous elimination we know that the inequality $\sum_{i=1}^{n} M_i \leq I$, can be be replaced with an equality without affecting the semi-definite program. So we can formulate the program as,

$$A = \begin{pmatrix} p_1\rho_1 & \\ & p_2\rho_2 \end{pmatrix} \tag{2.43}$$

$$X = \begin{pmatrix} M_1 & \\ & M_2 \end{pmatrix}. \tag{2.44}$$

To combine the two constraints $\sum_{i=1}^{n} M_i \leq I$ and $Tr[\rho_i M_j] = 0 \,\forall\, i \neq j$ we can define $\phi(X)$ as,

$$\phi(X) = \begin{pmatrix} M_1 + M_2 & & \\ & Tr[\rho_1 M_2] & \\ & & Tr[\rho_2 M_1] \end{pmatrix}, \tag{2.45}$$

and B as

$$B = \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.46}$$

Then using equation (2.30) and the same Y as in (2.32) to give us,

$$Y = \begin{pmatrix} N & & \\ & a_1 & \\ & & a_2 \end{pmatrix} \quad \text{and} \quad \phi^*(Y) = \begin{pmatrix} N + a_2\rho_2 & \\ & N + a_1\rho_1 \end{pmatrix}. \tag{2.47}$$

This gives us the connected pair of programs

Primal Problem

sup $\sum_{i=1}^{2} Tr[p_i\rho_i M_i]$.
Subject to $M_1 + M_2 \leq I$,
$Tr[\rho_1 M_2] = 0,\ Tr[\rho_2 M_1] = 0$
$M_1, M_2 \geq 0$.

Dual Problem

inf $Tr[N]$.
Subject to $N + a_1\rho_1 \geq \rho_2$,
$N + a_2\rho_2 \geq \rho_1$,
$a_i \in \mathbf{R}\ \ \forall i$,
$N \in$ Herm.

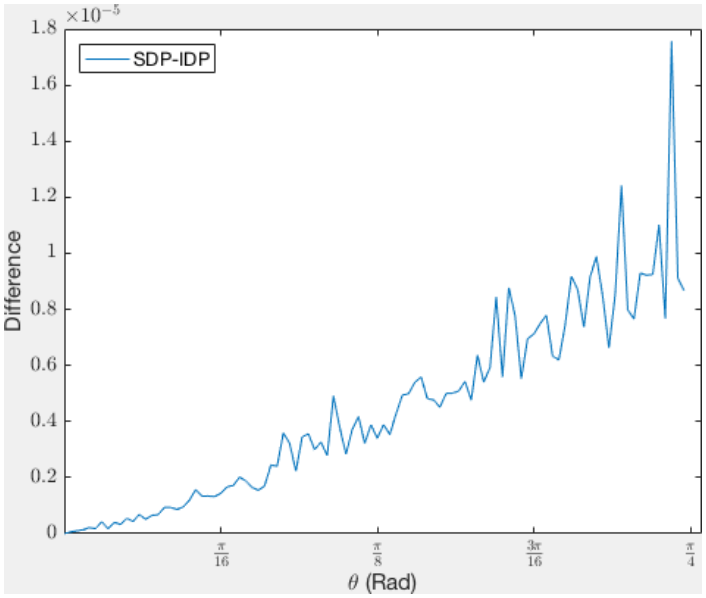


Figure 2.1: Difference between the SDP result for unambiguous discrimination between the two states $|+\theta\rangle$ and $|-\theta\rangle$, and the IDP limit of $1 - \cos(2\theta)$. Where the y axis represents the magnitude of the difference. As we can see the difference goes from 0 to $\sim 1x10^{-5}$. This difference is the duality gap. We are expecting a gap as the program doesn't have strong duality. It is hard to tell if the constant jitter is a result of the program or if it is due to numerical error. We discuss this more in the results section of this chapter.

To tidy up this piece for the future it would be a good idea to try and find an analytic solution the dual problem and comparing it to the IDP result to see the duality gap.

2.10 Implementation Of SDP

A popular method of solving these convex optimisation problems is to use the CVX package which utilises Matlab [38]. It allows users to formulate the constraints and objectives

in the Matlab language. Fitting the problem into Matlab syntax is fairly simple. In the SDP mode of the CVX package inequalities such as, $X >= Y$ become $X - Y ==$ hermitian_semidefinite(n), where n is the dimensions of the matrices X and Y. The $==$hermitian_semidefinite(n) syntax sets the requirement that $X - Y$ must be hermitian and positive semidefinite.

2.10.1 Two Qubits

The states $|\theta\rangle$ and $|-\theta\rangle$ are defined as:

$$\begin{aligned} |\theta\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle, \\ |-\theta\rangle &= \cos\theta|0\rangle - \sin\theta|1\rangle. \end{aligned} \tag{2.48}$$

In the two-qubit case the possible states are:

$$|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle. \tag{2.49}$$

In the computational basis these can be written as:

$$\begin{aligned} |\theta,\theta\rangle &= \cos^2\theta|00\rangle + \cos\theta\sin\theta|01\rangle + \cos\theta\sin\theta|10\rangle + \sin^2\theta|11\rangle, \\ |\theta,-\theta\rangle &= \cos^2\theta|00\rangle - \cos\theta\sin\theta|01\rangle + \cos\theta\sin\theta|10\rangle - \sin^2\theta|11\rangle, \\ |-\theta,\theta\rangle &= \cos^2\theta|00\rangle + \cos\theta\sin\theta|01\rangle - \cos\theta\sin\theta|10\rangle - \sin^2\theta|11\rangle, \\ |-\theta,-\theta\rangle &= \cos^2\theta|00\rangle - \cos\theta\sin\theta|01\rangle - \cos\theta\sin\theta|10\rangle + \sin^2\theta|11\rangle. \end{aligned} \tag{2.50}$$

Defining the basis vectors as

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.51}$$

we can now have the four states defined in as 4-component vectors, which can be input into Matlab. From this we produce a program that takes the states and starting from $\theta = 0$ we incrementally increase θ by some pre-defined step to some final angle whilst applying a semi-definite program for each angle.

2.10.2 Three Qubits

We now consider three qubits. each of which can be in the state $|\theta\rangle$ or $|-\theta\rangle$. The eight possible three-qubit states are then,

$$\begin{aligned} |\pm\theta,\pm\theta,\pm\theta\rangle = \cos^3\theta|000\rangle \pm \cos^2\theta\sin\theta|001\rangle \pm \cos^2\theta\sin\theta|010\rangle \pm \cos^2\theta\sin\theta|011\rangle \\ \pm \cos\theta\sin^2\theta|100\rangle \pm \cos\theta\sin^2\theta|101\rangle \pm \cos\theta\sin^2\theta|110\rangle \pm \sin^3\theta|111\rangle, \end{aligned} \tag{2.52}$$

	$\|000\rangle$	$\|001\rangle$	$\|010\rangle$	$\|011\rangle$	$\|100\rangle$	$\|101\rangle$	$\|110\rangle$	$\|111\rangle$
$\|\theta,\theta,\theta\rangle$	+	+	+	+	+	+	+	+
$\|\theta,\theta,-\theta\rangle$	+	-	+	+	-	-	+	-
$\|\theta,-\theta,\theta\rangle$	+	+	-	+	-	+	-	-
$\|-\theta,\theta,\theta\rangle$	+	+	+	-	+	-	-	-
$\|\theta,-\theta,-\theta\rangle$	+	-	-	+	+	-	-	+
$\|-\theta,\theta,-\theta\rangle$	+	-	+	-	-	+	-	+
$\|-\theta,-\theta,\theta\rangle$	+	+	-	-	-	-	+	+
$\|-\theta,-\theta,-\theta\rangle$	+	-	-	-	+	+	+	+

Table 2.5: Table showing the sign of the coefficients for the three qubit states with the magnitudes given in equation (2.52)

.

where the signs of the coefficients are given in table 2.5.

We assign the basis vectors in standard form as

$$|\vec{x}\rangle = \begin{pmatrix} 0 & 0 & \dots & 1 & \dots & 0 & 0 & 0 & 0 \end{pmatrix}^{\dagger}, \tag{2.53}$$

where the the 1 lies in the $x+1$ position with x being given as a binary number. It is $x+1$ as we begin counting from 0 instead of 1. For example

$$|011\rangle = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^{\dagger}, \tag{2.54}$$

as 011 in binary is 3 in base 10 and so the 1 is in the 4^{th} position. For two and three qubits we now have the states to be eliminated in a form Matlab can handle. Our desired output is a plot of success probability against θ. We will vary θ from 0 to $\pi/4$. $\pi/4$ is chosen as it is the point at which the states are orthogonal and then we know that individual measurement on each qubit will succeed with certainty.

2.10.3 SDP Code Example

Below is a section of the SDP code to find the bound for unambiguous elimination of one of the four two qubit states.

```
function y= Ustates2_1dual(rho1,rho2,rho3,rho4)
cvx_begin sdp quiet
cvx_solver sdpt3
cvx_precision best
variable N(4,4) hermitian
variable a1
variable a2
```

```
variable a3
variable a4
minimize trace(N)
subject to
a1*rho1+N-(rho1+rho2+rho3+rho4)>=0;
a2*rho2+N-(rho1+rho2+rho3+rho4)>=0;
a3*rho3+N-(rho1+rho2+rho3+rho4)>=0;
a4*rho4+N-(rho1+rho2+rho3+rho4)>=0;
N==hermitian_semidefinite(4);
cvx_end
y=cvx_optval
```

Listing 2.1: The SDP code for eliminating one of four two-qubit states unambiguously implementing the dual program defined in (2.34)

The code is generally split into sections with the top line from listing 2.1 defining the function that will be called in the program we stated earlier that runs through the states for varying value of $|\theta\rangle$. In the two-qubit case rho1, rho2, rho3 and rho4 will be $|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle$ respectively. The next three lines begin the CVX package, state the type of solver and precision of the CVX package. Then the variables $N, a1, a2, a3$ and $a4$ are defined. These are what the algorithm varies to optimise the aim. Then we have the problem to be optimised (minimize trace(N)), followed by the required conditions $N + a_i\rho_i \geq \sum_{j=1}^{n} \rho_j$ and $N \in$ Herm. Finally the CVX program is ended and the optimum value is the output of the function. Once we apply this function to the program that inputs the states with varying θ we will have a result for each value of θ ready to be plotted.

2.11 Results

In all the plots it is the dual program that is calculated by the code and plotted. This means upper limits are produced for all SDP plots. For the minimum error case the optimal solution to the dual is equivalent to the optimal solution to the primal problem. Whereas for the unambiguous measurements the results plotted are the optimal solution for the dual problem which is an upper bound as strong duality does not hold.

2.11.1 Results For Minimum-Error Elimination Measurements

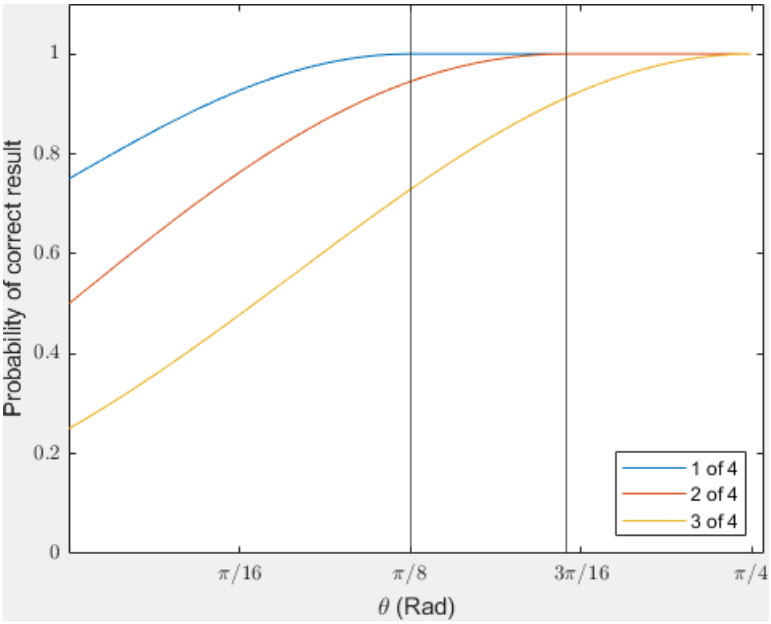


Figure 2.2: Convex optimisation results for minimum-error quantum state elimination on the two-qubit states. Eliminating one, two or three of the four possible options $|\pm\theta, \pm\theta\rangle$. *The vertical lines show the points at which the probability of success becomes one. To eliminate one of four this is at* $\pi/8$ *as expected as this is the PBR measurement, which saturates the bound from equation* (1.39). *To eliminate two of four states this is certainty at* $\theta \approxeq 0.571(rad) \approxeq 37.75°$.

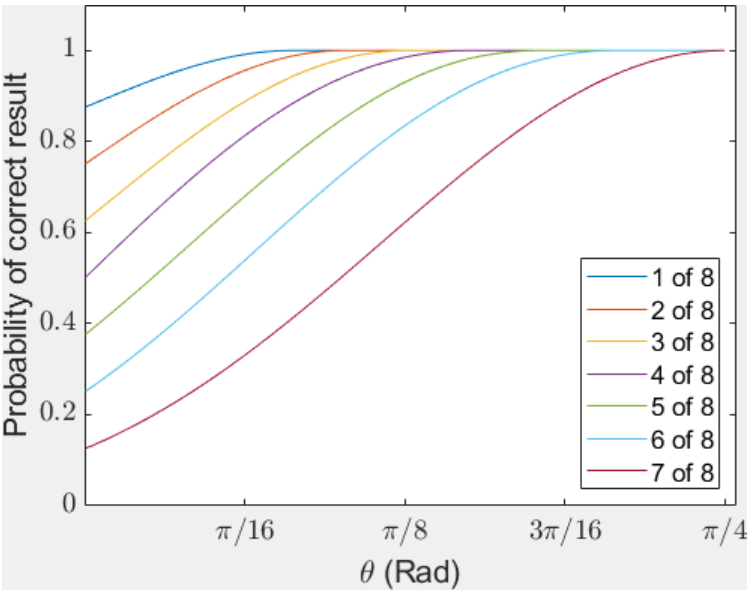
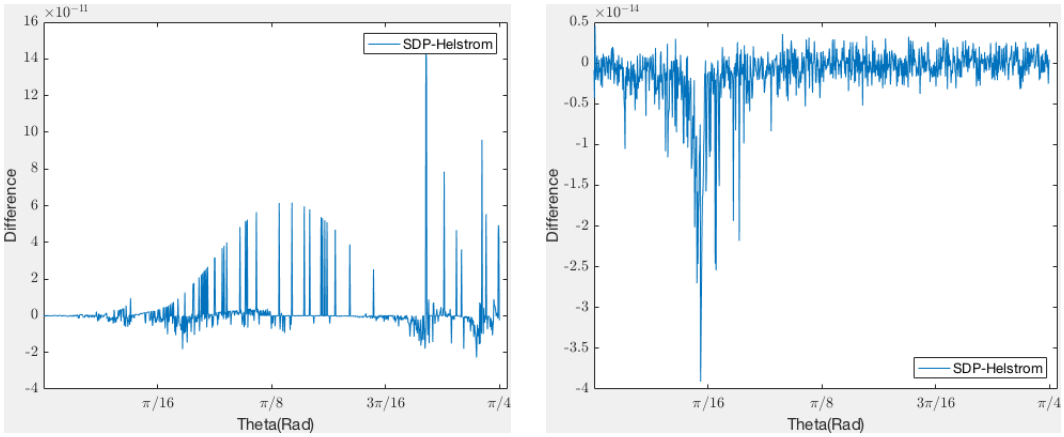


Figure 2.3: Convex optimisation results for minimum-error quantum state elimination on the three-qubit states. Eliminating from the eight possible options $|\pm\theta, \pm\theta, \pm\theta\rangle$.

As a quick check of the validity of the results from figures 2.2 and 2.3 we can compare parts to proven results we already know. For example when the states are the same ($\theta = 0$)



(a) Two Qubits. *(b) Three Qubits.*

Figure 2.4: Plots showing the difference between the success probabilities obtained from SDP and the Helstrom bound, where the y axis is the value of that difference. The difference is mostly around 0 with some values of up to $4x10^{-14}$*. This small amount is most likely due to computational error. We discuss the fluctuations in the upcoming section, duality gap or numerical errors.*

then the probability of successful elimination should just be that of randomly guessing. This is $p = 1 - m/2^N$ where m is the number of states we aim to eliminate and N is the number of qubits. We can easily see in the two-qubit case we have this for $m = 1, 2, 3,\ p = 0.75, 0.5, 0.25$ as expected. The same applies in the three-qubit case in that we have a sensible solution for $\theta = 0$. Furthermore as eliminating all but one state is equivalent to quantum state discrimination we can check the validity of the three out of four and seven out of eight by comparing it to the Helstrom bound [35] for consecutive measurements on the individual qubits. This is because Helstrom showed that for state discrimination the probability of successfully guessing an outcome could not exceed the Helstrom bound given as,

$$P_{err} = \frac{1}{2}\left(1 - \sqrt{1 - 4p_0p_1|\langle\psi_0|\psi_1\rangle|^2}\right), \tag{2.55}$$

where p_0 and p_1 are the prior probabilities of states $|\psi_0\rangle$ and $|\psi_1\rangle$ respectively. In our case we have $p_0 = p_1 = 1/2$ and the overlap $|\langle\psi_0|\psi_1\rangle| = \cos(2\theta)$. As we are interested in success probabilities $(1 - P_{err})$ and also looking at a varying number of qubits we have

$$P_{suc} = \left[\frac{1}{2}\left(1 + \sqrt{1 - cos^2(2\theta)}\right)\right]^N, \tag{2.56}$$

where N is the number of qubits.

Duality Gap or Numerical Errors?

Figures 2.4 seem to suggest that SDP is accurate with respect to the Helstrom measurement. The errors are very small and due to the irregularity of them this suggest that it is

most likely numerical error that has occurred from the computation process. This belief comes from the magnitudes of the errors involved and the form of the difference being very irregular.

As you can see from figures 2.4 the largest difference is of magnitude $\approx 10^{-10}$, which is extremely small, yet it is the form of the difference that I believe gives the most evidence it is likely numerical error. Firstly if it were weak duality we would have an upper bound on the results and therefore the difference would only be positive yet we see in both cases there are negative spikes. Also the difference is fairly irregular with just spikes here and there and no obvious function. Whereas if you compare this to figure 2.1 in which we know we have a duality gap, there is a more consistent gap between the known optimal solution and the result from SDP and the difference is always positive.

2.11.2 Results For Unambiguous State Elimination Measurements

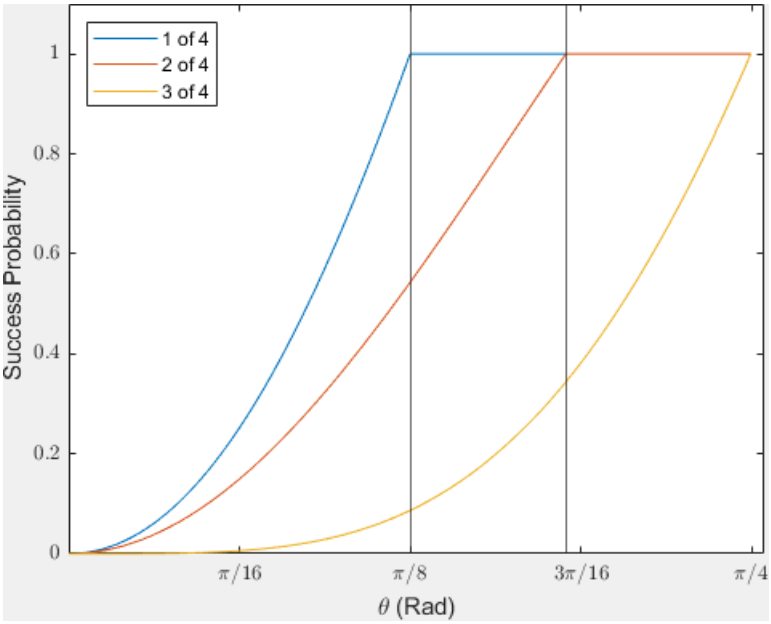


Figure 2.5: Convex optimisation results for unambiguous quantum state elimination on the set of two-qubit states. The plots show the success probabilities of eliminating one, two or three of the four possible options $|\pm\theta, \pm\theta\rangle$. The vertical lines show the point at which the unambiguous measurement succeeds every time. Comparing with figure 2.2 we see this occurs at the same angles at expected.

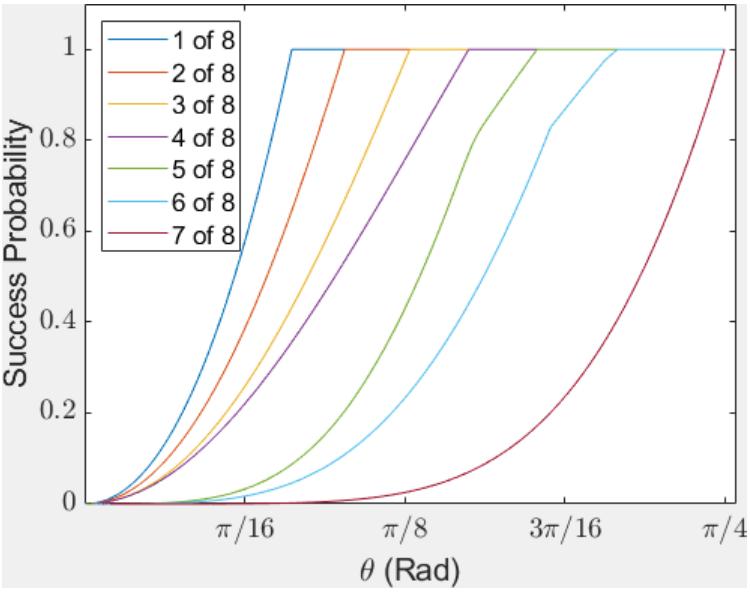


Figure 2.6: Convex optimisation results for unambiguous quantum state elimination on three-qubit states that eliminates one to seven of the eight possible options $|\pm\theta, \pm\theta, \pm\theta\rangle$.

For the results on unambiguous measurements, we can make similar checks as we made for the minimum-error measurements to check the validity of the results. For an unambiguous measurement, the success probability should be zero when $\theta = 0$, as it is impossible to unambiguously distinguish between two identical states. Unambiguous state discrimination is equivalent to eliminating all but one state and there is a proof using no-signalling [39] that shows individual unambiguous measurements on each qubit is optimal for discriminating a sequence of states. Therefore we can use the IDP bound as a limit for eliminating all but one state as we did with the Helstrom bound for minimum error. Finally, when the success probability of unambiguous state discrimination is 1, then the probability for a minimum-error measurement is also 1, and vice versa, therefore we can compare the value of θ for those two cases.

There is a caveat in the unambiguous case that we do not have strong duality and therefore the results shown in figures (2.5) and (2.6) are just upper bounds to the optimal success probabilities. This is until $p = 1$ then the unambiguous program becomes equivalent to the minimum error one and at the point we have strong duality.

The IDP limit for unambiguous discrimination between two states with equal a priori probability states that the probability of an inconclusive result is given by the overlap of the two states. Therefore the probability of success is given by,

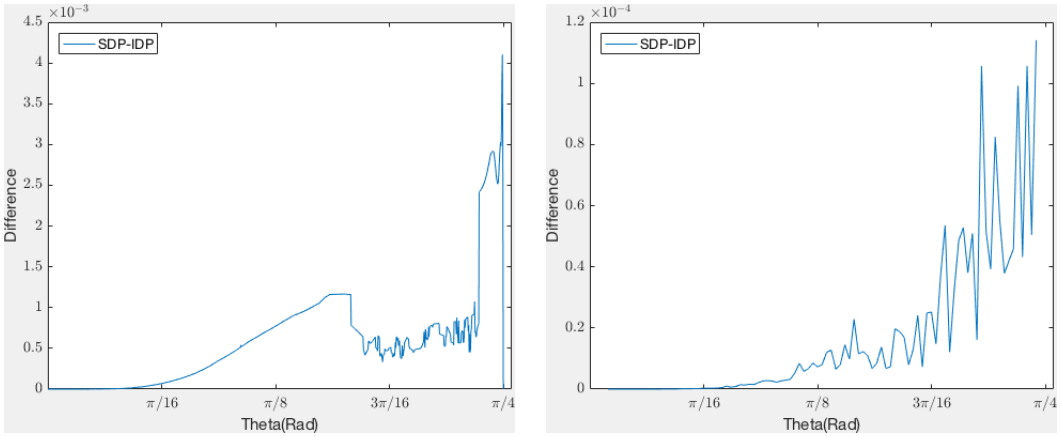
$$1 - |\langle\psi_0|\psi_1\rangle| = 1 - |\langle\theta|{-\theta}\rangle| = 1 - \cos(2\theta). \tag{2.57}$$

For each qubit we require all measurements to be successful therefore the optimal success

probability for eliminating all but one state is given by,

$$p_s = (1 - \cos(2\theta))^N, \tag{2.58}$$

where N is the number of qubits. We compare this to the SDP results for eliminating $2^N - 1$ of the 2^N possible state preparations for $N = 2$ and $N = 3$.



(a) Two Qubits. *(b) Three Qubits.*

Figure 2.7: The difference between the success probabilities obtained from SDP and the IDP bound to the power of the number of qubits, where the y axis represents the magnitude of that difference. The difference in both cases is gradually increasing with θ and reaches a maximum of approximately $1\text{x}10^{-4}$. As we can see the magnitude of this difference is much larger than in figure 2.4 with an order of 10^{-4} instead of 10^{-10}. This is then likely to be the duality gap.This idea is backed up by the difference being of a comparable magnitude to that from USD.

.

In figure 2.6 the curves that represent the results for eliminating 5 and 6 out of the 8 states have an interesting looking kink in them around the point the success probability is 0.8. The bends seem to be instantaneous and even with more data points around the bend it doesn't smooth out. We are unsure what causes these as we haven't looked into the three qubit measurement thoroughly.

2.11.3 Minimum Angle For Conclusive Elimination

Here I present the minimum angle θ required to eliminate certain numbers of states in the two- and three-qubit case unambiguously with a success probability of 1.

Two Qubits

Number of States Eliminated	Minimum Angle
1	$22.5°$
2	$32.765°$
3	$45°$

Table 2.6: Table showing the minimum angle required to eliminate a certain number of states 100% of the time for two qubits. To eliminate one state with certainty we see the angle is $22.5°$ as expected as this is the PBR measurement. To eliminate 3 with certainty we require the states to be orthogonal as this is just state discrimination.

Three Qubits

Number of States Eliminated	Minimum Angle
1	$14.570°$
2	$18.284°$
3	$22.725°$
4	$27.014°$
5	$32.012°$
6	$37.467°$
7	$45°$

Table 2.7: Table showing the minimum angle required to eliminate a certain number of states 100% of the time for two qubits.

For both two and three qubits as θ increases the number of states that can be eliminated increases as we would expect. At the point $\theta = 45°$ perfect state discrimination is available and hence it is possible to eliminate all but one state. The semidefinite program in (2.34) can be produced for any number of qubits. So far I have just considered two and three qubits. The more qubits there are the longer the code becomes as there are more possible ways to eliminate different numbers of states. The number of constraints is given by

$$\sum_{m=1}^{n} \binom{n}{m}. \tag{2.59}$$

2.12 Conclusion

In this chapter I have introduced semi-definite programming in order to illustrate its potential for numerical work in quantum theory. It proved a very useful tool at finding the optimal success probabilities for quantum state elimination measurements. It would be interesting to investigate the duality gap for unambiguous measurements a bit more as it seems there aren't too many cases in quantum information where strong duality doesn't hold. In the next chapter we will compare the results produced in this chapter to the success probabilities of measurements we have obtained analytically. Therefore being able to find not just the optimal success probabilities but also the measurements that can reach those probabilities.

Chapter 3

Analytic Methods

3.1 Unambiguously Eliminating One State In The Two-Qubit Case

We started looking at the previous elimination measurements including the PBR measurement as shown in (1.33) and attempted to extend this for generalised angles and into the cases when the measurement does not succeed with certainty every time. To unambiguously eliminate one state in the two-qubit case we have the initial four two-qubit possibilities $|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle$, as given in (1.46), and the aim is to unambiguously eliminate one of these four possibilities.

3.1.1 Generalised PBR Measurement

We started by taking a similar approach as that done in the PBR measurement, by constructing the orthogonal state out of an equal proportion of two orthogonal states that each had a single qubit orthogonal and a single qubit the same. For example for $|00\rangle$ PBR used the orthogonal state

$$|\psi_{\overline{00}}\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle), \tag{3.1}$$

as $|1\rangle$ is orthogonal to $|0\rangle$. Therefore with our initial states the four orthogonal states are given by

$$\begin{aligned}
|\psi_{\overline{\theta,\theta}}\rangle &= \frac{1}{\sqrt{2}}(|\overline{\theta},\theta\rangle + |\theta,\overline{\theta}\rangle),\\
|\psi_{\overline{\theta,-\theta}}\rangle &= \frac{1}{\sqrt{2}}(|\overline{\theta},-\theta\rangle + |\theta,\overline{-\theta}\rangle),\\
|\psi_{\overline{-\theta,\theta}}\rangle &= \frac{1}{\sqrt{2}}(|\overline{-\theta},\theta\rangle + |-\theta,\overline{\theta}\rangle),\\
|\psi_{\overline{-\theta,-\theta}}\rangle &= \frac{1}{\sqrt{2}}(|-\theta,\overline{-\theta}\rangle + |-\theta,\overline{-\theta}\rangle),
\end{aligned} \tag{3.2}$$

where each state is orthogonal to one of the generalised states in (1.44), and where

$$|\overline{\theta}\rangle = \sin\theta|0\rangle - \cos\theta|1\rangle,$$
$$|\overline{-\theta}\rangle = \sin\theta|0\rangle + \cos\theta|1\rangle, \tag{3.3}$$

are orthogonal to the individual qubit states $|\theta\rangle$ and $|-\theta\rangle$. We called this the generalised PBR measurement as it very similar to the original PBR measurement, differing from using a range of initial states. In the case where $\theta = 22.5^\circ$ we have the same overlaps as in the original PBR measurement and the method will unambiguously eliminate one of the four states every single time.

We will now construct a measurement where some of the measurement operators are proportional to projectors onto the states in (3.2). Each corresponding outcome eliminates one two-qubit state. In addition, there might also be a "failure" measurement operator required to complete the measurement.

In any quantum measurement, the measurement operators should sum to the identity. If we add up projectors onto the four states in (3.2), we obtain

$$\begin{aligned} M =& |\psi_{\overline{\theta},\overline{\theta}}\rangle\langle\psi_{\overline{\theta},\overline{\theta}}| + |\psi_{\overline{\theta},\overline{-\theta}}\rangle\langle\psi_{\overline{\theta},\overline{-\theta}}| + |\psi_{\overline{-\theta},\overline{\theta}}\rangle\langle\psi_{\overline{-\theta},\overline{\theta}}| + |\psi_{\overline{-\theta},\overline{-\theta}}\rangle\langle\psi_{\overline{-\theta},\overline{-\theta}}| \\ =& \begin{pmatrix} 8\sin^2\theta\cos^2\theta & 0 & 0 & 0 \\ 0 & 8\cos^4\theta - 8\cos^2\theta + 2 & 0 & 0 \\ 0 & 0 & 8\cos^4\theta - 8\cos^2\theta + 2 & 0 \\ 0 & 0 & 0 & 8\sin^2\theta\cos^2\theta \end{pmatrix} \end{aligned} \tag{3.4}$$

This matrix happens to be diagonal in the basis we are using, and its eigenvalues are therefore the diagonal elements, which by construction are all positive, but can be larger than 1. In order to construct a valid quantum measurement, we can multiply the projectors onto the states in (3.2) with a factor which is less than one. In the regions $0 \leq \theta \leq \frac{\pi}{8}$ and $\frac{3\pi}{8} \leq \theta \leq \frac{\pi}{2}$, the term $8\cos^4\theta - 8\cos^2\theta + 2$ is the largest and so we divide (3.4) by this factor, then rearrange to give the matrix

$$\tilde{M}_1 = \begin{pmatrix} \tan^2(2\theta) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \tan^2(2\theta) \end{pmatrix}. \tag{3.5}$$

In the region $\frac{\pi}{8} \leq \theta \leq \frac{3\pi}{8}$, the $8\sin^2\theta\cos^2\theta$ term is the largest so again dividing (3.4) by this factor we obtain the matrix

$$\tilde{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cot^2(2\theta) & 0 & 0 \\ 0 & 0 & \cot^2(2\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.6}$$

$\tilde{M}_1$ and $\tilde{M}_2$ are the sum of the measurement operators corresponding to the "success" outcomes for each region of θ, where "success" means that one two-qubit state has been eliminated. The probability of success is just given by

$$P_s = Tr\left(\tilde{M}\rho\right). \tag{3.7}$$

$\tilde{M}$ is the sum of the measurement operators and ρ is the measured state that can be given by

$$\begin{aligned}\rho &= \cos^4\theta|00\rangle\langle 00| + \sin^4\theta|11\rangle\langle 11| + \cos^2\theta\sin^2\theta(|01\rangle\langle 01| + |10\rangle\langle 10|) \\ &= \begin{pmatrix} \cos^4\theta & 0 & 0 & 0 \\ 0 & \cos^2\theta\sin^2\theta & 0 & 0 \\ 0 & 0 & \cos^2\theta\sin^2\theta & 0 \\ 0 & 0 & 0 & \sin^4\theta \end{pmatrix},\end{aligned} \tag{3.8}$$

assuming equal a priori probabilities. The success probability of the measurement is therefore

$$p_{s1} = \tan^2(2\theta)(\cos^4\theta + \sin^4\theta) + 2\cos^2\theta\sin^2\theta, \tag{3.9}$$

for $0 \leq \theta \leq \frac{\pi}{8}$ and $\frac{3\pi}{8} \leq \theta \leq \frac{\pi}{2}$ and

$$p_{s2} = \cos^4(\theta) + \sin^4(\theta) + 2\cot^4(2\theta)\cos^2(\theta)\sin^2(\theta) = \frac{\cos(4\theta)}{2} + 1, \tag{3.10}$$

for $\frac{\pi}{8} \leq \theta \leq \frac{3\pi}{8}$.

The success probability for this generalised PBR measurement is shown in figure 3.1. From this we can see the measurement gives conclusive results with probability 1 for $\theta = \pi/8$, this is as expected since in this limit the states in (1.44) now have the same overlap as the states $|0\rangle$ and $|+\rangle$ from the PBR measurement shown in chapter 1.3. For the region $\pi/8 \geq \theta \geq \pi/4$ we know that the optimal measurement will give a probability of one to eliminate a state. This comes from the fact that in the PBR paper [1] it was proved the number of qubits required to eliminate a single state with 100% success for a given angle was

$$2^{\frac{1}{n}} - 1 \leq \tan\left(\frac{\theta}{2}\right). \tag{3.11}$$

At the point $\theta = \frac{\pi}{8}$ then we require $n \geq 2$. Also intuitively we have a measurement that succeeds 100% of the time for $\pi/8$. As the angle θ increases the states become more distinguishable and so it makes sense that the success probability should not decrease. From figure (3.1) we see that for $\theta > \pi/8$ we don't have a probability of one so know our measurement is not optimal in this realm. From figure (3.2) we see the measurement does not reach the bound given by semi-definite programming. The difference between the SDP bound and the generalised measurement is shown in figure 3.3. From this you

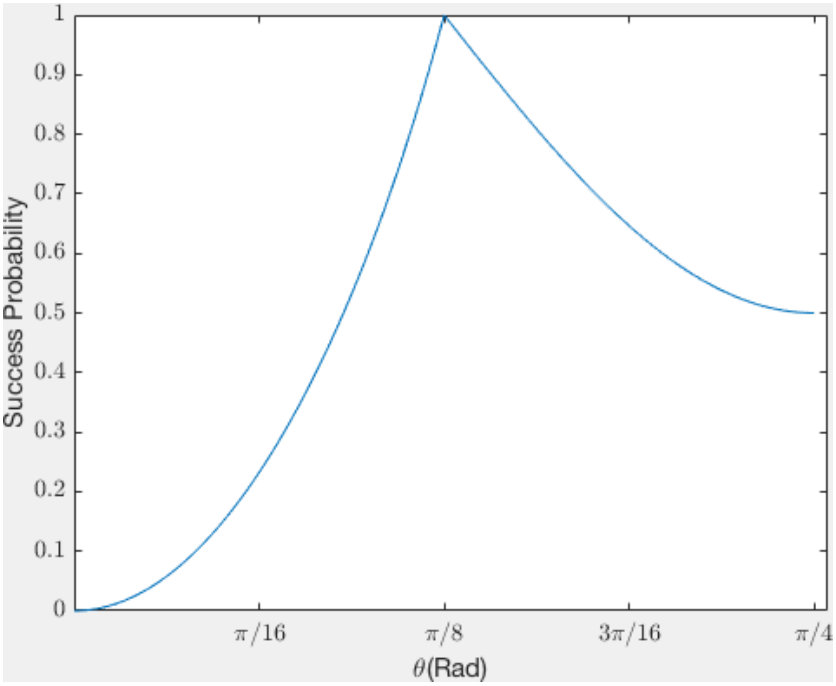


Figure 3.1: The probability of a successful measurement using the generalised PBR measurement. The success probability is equal to zero for $\theta = 0$ and equal to 1 for $\theta = \pi/8$, as expected. For $\theta > \pi/8$ we know it is sub-optimal as the success probability should be one.

can see the difference starts at 0 because both measurements give a probability of 0 for unambiguous discrimination between identical states as expected. Then generally as θ increases so does the difference between the SDP bound and the generalised result. The difference then falls to 0 as both results give a perfect success probability for $\theta = \pi/8$.

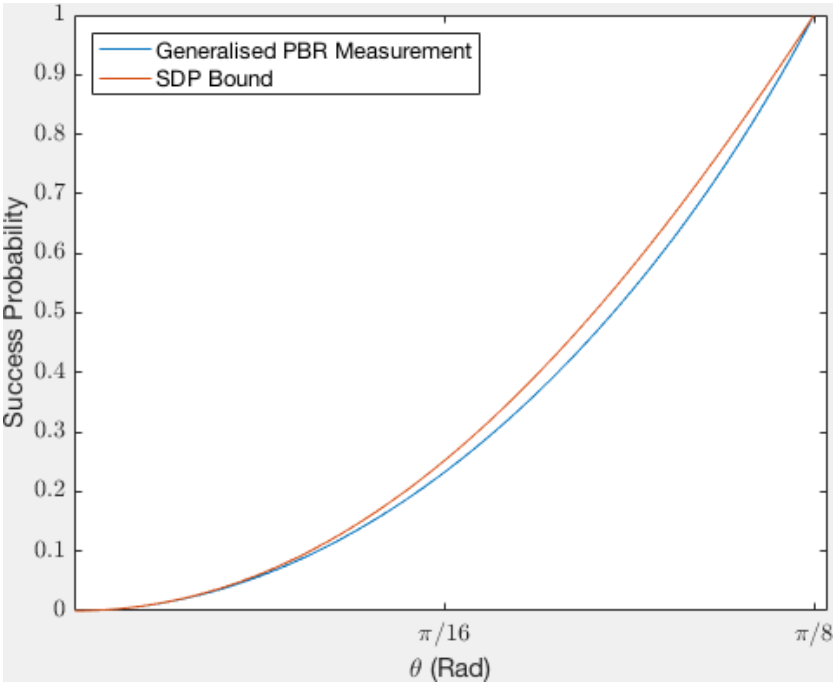


Figure 3.2: The comparison of success probability from the generalised PBR method and the SDP. The bound on the success probability using SDP is higher than the success probability of the generalised PBR measurement for all θ excluding $\theta = 0$ and $\theta = \pi/8$.

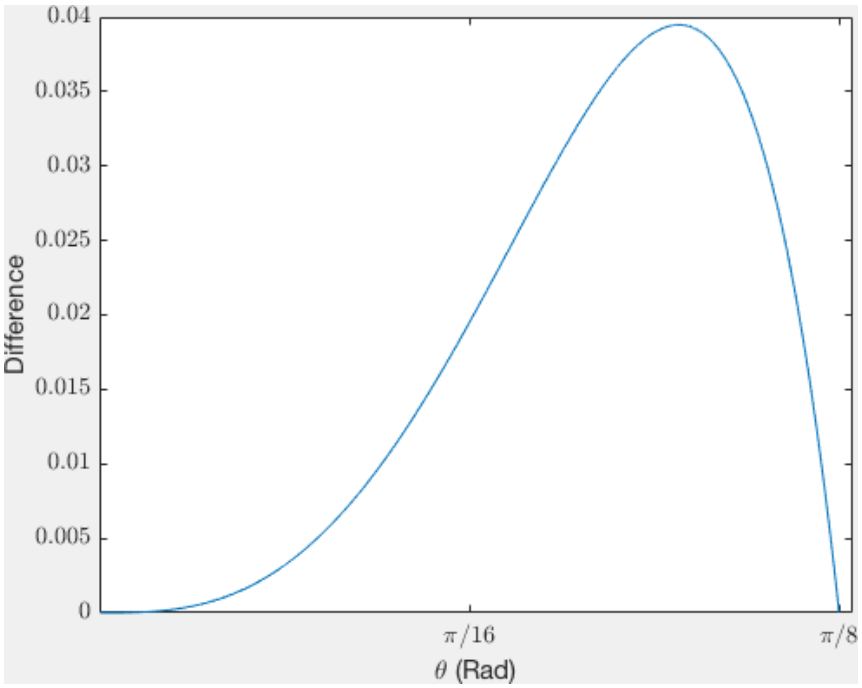


Figure 3.3: The difference between the probability of a successful outcome obtained from the semi-definite program and the generalised PBR method. The maximum difference is about 0.04, and the difference is generally larger for higher θ.

3.1.2 Nearest 22.5° Approach

Since the 'generalised PBR' Ansatz did not give the optimal measurement, we will attempt to construct another possible measurement to see if we can find the optimal result. The 'nearest 22.5' method is based on finding the state that is orthogonal to $|\theta\theta\rangle$, and also closest to the measurement states from (3.2) with the angle at $\theta = 22.5°$. At $\theta = 22.5°$ (3.2) are equivalent to original PBR measurement states aiming to eliminate with the two qubits $|\theta\rangle, |-\theta\rangle$ instead of $|0\rangle$ and $|+\rangle$. We label these measurement states as $|\sigma_{PBR++}\rangle$ to eliminate the $|\theta,\theta\rangle$, $|\sigma_{PBR-+}\rangle$ to eliminate $|-\theta,\theta\rangle$ and so on. Then the nearest orthogonal state can be produced by projecting the $|\sigma_{PBR++}\rangle$ state onto $|\theta,\theta\rangle$ in the form

$$\begin{aligned}
|\psi_{\overline{\theta,\theta}}\rangle &= |\sigma_{PBR++}\rangle - \langle\sigma_{PBR++}|\theta,\theta\rangle|\theta,\theta\rangle,\\
|\psi_{\overline{\theta,-\theta}}\rangle &= |\sigma_{PBR+-}\rangle - \langle\sigma_{PBR+-}|\theta,-\theta\rangle|\theta,-\theta\rangle,\\
|\psi_{\overline{-\theta,\theta}}\rangle &= |\sigma_{PBR-+}\rangle - \langle\sigma_{PBR-+}|-\theta,\theta\rangle|-\theta,\theta\rangle,\\
|\psi_{\overline{-\theta,-\theta}}\rangle &= |\sigma_{PBR--}\rangle - \langle\sigma_{PBR--}|-\theta,-\theta\rangle|-\theta,-\theta\rangle.
\end{aligned} \tag{3.12}$$

At this stage these states aren't normalised but as we will be multiplying the projectors onto these states with some factor in the future we will satisfy the normality requirements at that stage. If we sum up the projectors onto the states from (3.12) with the same weighting then we obtain

$$M = |\psi_{\theta,\theta}\rangle\langle\psi_{\theta,\theta}| + |\psi_{\theta,-\theta}\rangle\langle\psi_{\theta,-\theta}| + |\psi_{-\theta,\theta}\rangle\langle\psi_{-\theta,\theta}| + |\psi_{-\theta,-\theta}\rangle\langle\psi_{-\theta,-\theta}|, \tag{3.13}$$

$$M = \begin{pmatrix} M_{11} & 0 & 0 & 0\\ 0 & M_{22} & 0 & 0\\ 0 & 0 & M_{33} & 0\\ 0 & 0 & 0 & M_{44} \end{pmatrix}, \tag{3.14}$$

where

$$\begin{aligned}
M_{11} &= (\cos^2\theta - 2\cos^4\theta + 2\cos^3\theta\sin\theta + 1)^2,\\
M_{22} &= M_{33} = \frac{1}{16}(\sin(8\theta) + 2(3\sqrt{2}\sin\left(4\theta + \frac{\pi}{4}\right)) + 10),\\
M_{44} &= 4(\sin^2\theta)\left(\left(\cos\theta\sin\theta - \frac{\cos^2\theta}{2} + \frac{\sin^2\theta}{2}\right) - \frac{1}{2}\right)^2.
\end{aligned} \tag{3.15}$$

Comparing the magnitudes of these diagonal elements for varying θ we can find the largest diagonal element and then divide by this factor so the eigenvalues of M don't exceed one. Figure (3.4) shows the values of each eigenvalue with respect to θ, so that we can see which is largest for what regions. As we can see in the region $(0 \leq \theta \leq \frac{\pi}{8})$ M_{22} and M_{33} are largest and in the region $(\frac{\pi}{8} \leq \theta \leq \frac{\pi}{4})$, M_{11} is largest. Using this information we can derive the optimal M that satisfies the condition of the eigenvalues being between 0 and

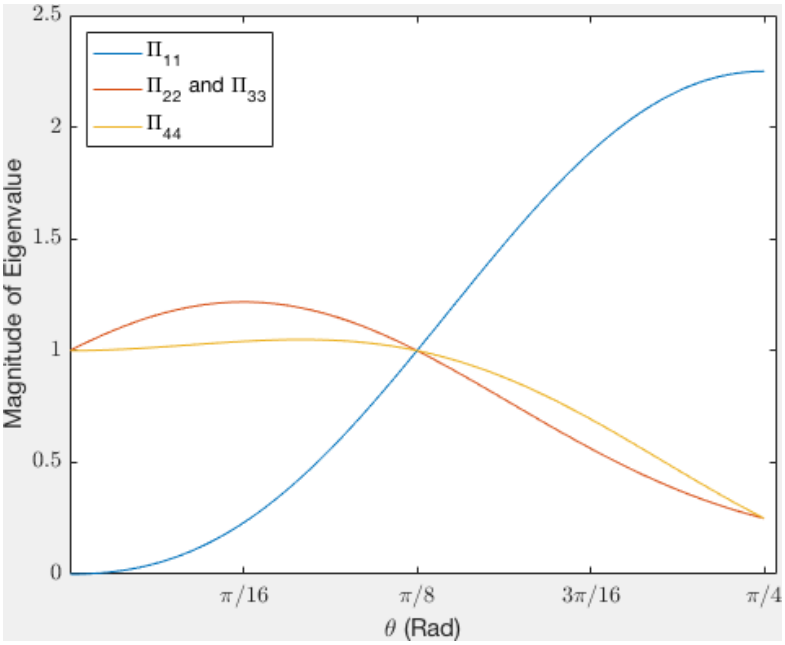


Figure 3.4: The eigenvalues of (3.14) *are plotted against* θ *. We observe that for* $0 \leq \theta \leq \pi/8$ *then the (2,2) and (3,3) elements are the largest diagonal elements and so we will divide* M *from* (3.14) *by these in that range so that the eigenvalues don't exceed 1. For* $\pi/8 \leq \theta \leq \pi/4$ *the (1,1) term is largest and so will be used reduce the size of the diagonal elements for that range.*

1 for this approach.

The new scaled sum of operators $\tilde{M}_{(0\leq\theta\leq\frac{\pi}{8})}$ and $\tilde{M}_{(\frac{\pi}{8}\leq\theta\leq\frac{\pi}{4})}$ for those specific regions are given by

$$\begin{aligned}\tilde{M}_{(0\leq\theta\leq\frac{\pi}{8})} &= \frac{M}{\frac{1}{16}(\sin(8\theta)+2(3\sqrt{2}\sin(4\theta+\frac{\pi}{4}))+10)} \\ \tilde{M}_{(\frac{\pi}{8}\leq\theta\leq\frac{\pi}{4})} &= \frac{M}{(\cos^2\theta-2\cos^4\theta+2\cos^3\theta\sin\theta+1)^2}.\end{aligned} \tag{3.16}$$

The success probability is then given by $Tr(\tilde{M}\rho)$ where ρ is given in equation (3.8). Figures 3.5, 3.6 and 3.7 show that the result is again non-optimal for $\theta > \pi/8$ and again still does not reach the SDP bound. It does improve upon the 'generalised PBR' measurement though.

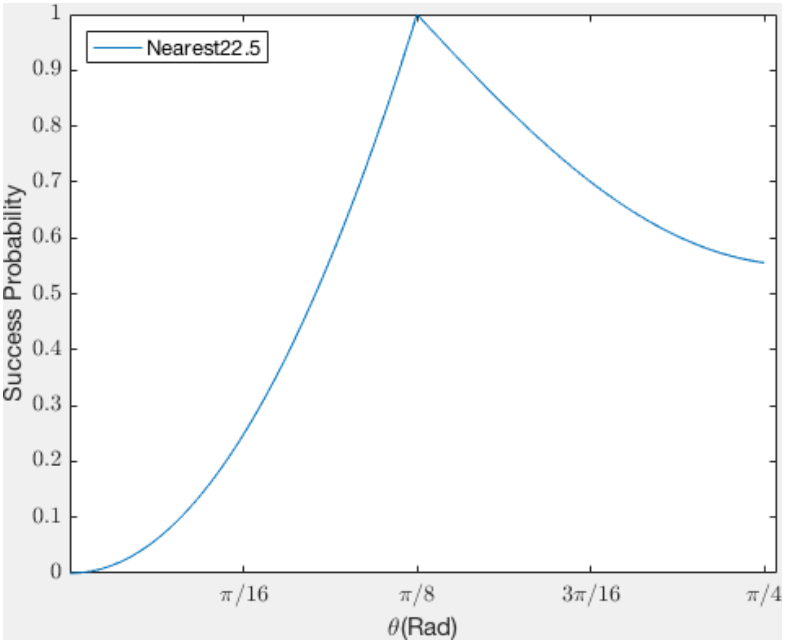


Figure 3.5: The success probability using the Nearest 22.5 approach. It has expected results for $\theta = 0$ *and* $\theta = \pi/8$ *. For* $\theta > \pi/8$ *again we know it is sub-optimal as the success probability should be one.*

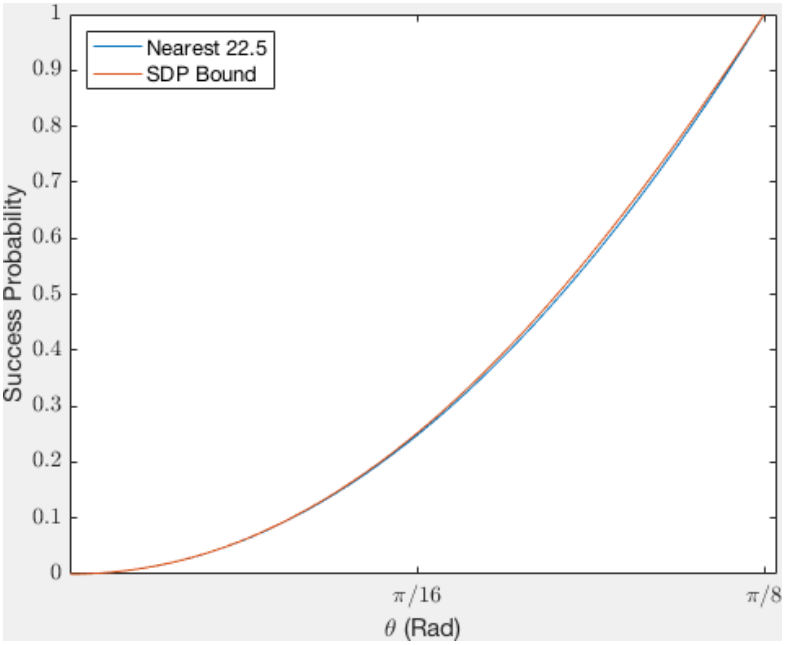


Figure 3.6: The comparison of success probability for the Nearest 22.5 measurement and the bound obtained using SDP. The bound from SDP is larger than the success probability of the Nearest 22.5 approach for all θ *except for* $\theta = 0$ *and* $\theta = \pi/8$ *when they are equal.*

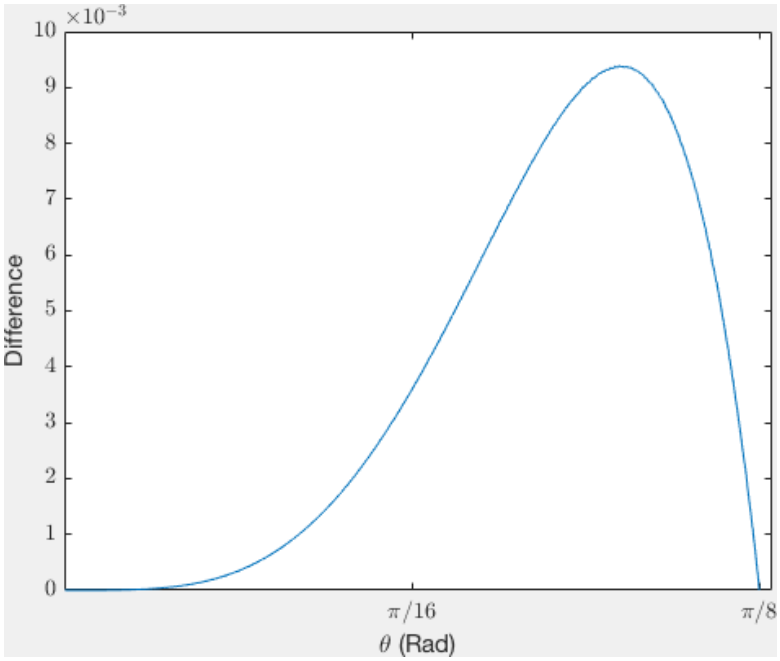


Figure 3.7: The difference between the bound on the probability of success given by SDP and the success probability for the Nearest 22.5 method. The y axis gives the magnitude of the difference between the two probabilities where the result from SDP is the higher value. As we can see the maximum difference is about 0.01 and the difference in general is larger for larger θ. Looking figure 3.3 we see that the maximum difference there is roughly 0.04 and in this figure it is 0.001 so the 'nearest 22.5' measurement performs better than the 'generalised PBR measurement'.

3.1.3 Derivation Using Group Theory

This approach was composed during work with Mark Hillery and turned out to produce the optimal measurement for $0 \leq \theta \leq 22.5°(\pi/8\ rad)$. In general a projector is

$$P = |X\rangle\langle X|, \tag{3.17}$$

where $|X\rangle$ is an unnormalised pure state,

$$|X\rangle = \sum_{j,k=0,1} c_{jk}|j\rangle \otimes |k\rangle. \tag{3.18}$$

For unambiguous elimination of $|\theta, \theta\rangle$ we require that $\langle\theta\theta|X\rangle = 0$. This leads to the requirement

$$c_{00}\cos^2\theta + (c_{01} + c_{10})\sin\theta\cos\theta + c_{11}\sin^2\theta = 0. \tag{3.19}$$

The group theory style of the approach comes about from studying the transformations of each two qubit state into another. The measurement for eliminating the other three states can then also be obtained by using the transformations between the states as will be shown in the upcoming work. First of all the transform from $|+\theta\rangle$ to $|-\theta\rangle$ is given by the

unitary $U = |0\rangle\langle 0| - |1\rangle\langle 1|$. This can then be applied to either the first, second or both of the qubits to get any of the four states from $|\theta, \theta\rangle$. This unitary performs the transform $U|\theta\rangle = |-\theta\rangle$ as can be seen by

$$\begin{aligned} U|\theta\rangle &= (|0\rangle\langle 0| - |1\rangle\langle 1|)(\cos\theta|0\rangle + \sin\theta|1\rangle), \\ &= \cos\theta|0\rangle - \sin\theta|1\rangle, \\ &= |-\theta\rangle. \end{aligned} \tag{3.20}$$

Then all four states two-qubit states, $|\theta, \theta\rangle, |\theta, -\theta\rangle, |-\theta, \theta\rangle, |-\theta-, \theta\rangle$, can be given by the following transformations upon the state $|\theta, \theta\rangle$,

$$\{I_1 \otimes I_2\,,\, I_1 \otimes U_2\,,\, U_1 \otimes I_2\,,\, U_1 \otimes U_2\}, \tag{3.21}$$

respectively. With the same transforms we can obtain all four of the measurement operators from $\Pi_{\overline{\theta,\theta}}$ in the following manner,

$$\begin{aligned} \Pi_{\overline{-\theta,\theta}} &= (U_1 \otimes I_2)\Pi_{\overline{\theta,\theta}}(U_1^\dagger \otimes I_2), \\ \Pi_{\overline{\theta,-\theta}} &= (I_1 \otimes U_2)\Pi_{\overline{\theta,\theta}}(I_1 \otimes U_2^\dagger), \\ \Pi_{\overline{-\theta,-\theta}} &= (U_1 \otimes U_2)\Pi_{\overline{\theta,\theta}}(U_1^\dagger \otimes U_2^\dagger). \end{aligned} \tag{3.22}$$

The failure operator that completes the measurement is given by

$$\Pi_{f,1} = I - \sum_{j,k=\pm\theta} \Pi_{\overline{(jk)}} = I - 4 \sum_{j,k=0,1} |c_{jk}|^2 |j\rangle\langle j| \otimes |k\rangle\langle k|. \tag{3.23}$$

Since $\Pi_{f,1} \geq 0$, it must hold that $|c_{jk}| \leq 1/2$. If the states are equally likely, then the failure probability is given by

$$\begin{aligned} p_{f,1} =& \frac{1}{4} \sum_{j,k=\pm\theta} \langle j, k|\Pi_f|j, k\rangle \\ =& 1 - 4[|c_{00}|^2 \cos^4\theta + (|c_{01}|^2 + |c_{10}|^2)\sin^2\theta\cos^2\theta + |c_{11}|^2 \sin^4\theta]. \end{aligned} \tag{3.24}$$

To minimise $p_{f,1}$ we need to maximise

$$|c_{00}|^2 \cos^4\theta + (|c_{01}|^2 + |c_{10}|^2)\sin^2\theta\cos^2\theta + |c_{11}|^2 \sin^4\theta. \tag{3.25}$$

For the range $0 \leq \theta \leq 22.5^\circ$ we have,

$$\cos^4\theta > \cos^2\theta \sin^2\theta > \sin^4\theta, \tag{3.26}$$

so we should make $|c_{00}|$ as large as possible. This can be done by making $c_{01} = c_{10} = c_{11} = -1/2$. This is because $-1/2$ is the smallest value the coefficients can take and considering the requirement in (3.19) the smallest values of c_{01}, c_{10} and c_{11} will lead to

the largest for c_{00} as for $\theta \leq 22.5°$, $\cos\theta \geq 0$ and $\sin\theta \geq 0$. Then from the requirement in (3.19) c_{00} can be calculated from

$$\begin{aligned} c_{00}\cos^2\theta - \sin\theta\cos\theta - \frac{1}{2}\sin^2\theta &= 0, \\ c_{00} &= \frac{\sin\theta}{\cos\theta} + \frac{1}{2}\frac{\sin^2\theta}{\cos^2\theta}, \\ c_{00} &= \tan\theta(1 + \frac{1}{2}\tan\theta). \end{aligned} \tag{3.27}$$

c_{00} is always less than or equal to $1/2$ as for $\theta = 22.5°$, $c_{00} = 1/2$ and isn't higher for any other value of theta in our range. The failure probability from equation (3.24) is then

$$\begin{aligned} p_{f,1} &= 1 - 4[(\tan\theta + \frac{1}{2}\tan^2\theta)^2\cos^4\theta + \frac{1}{2}\sin^2\theta\cos^2\theta + \frac{1}{4}\sin^4\theta] \\ &= 1 - 4\left(\frac{3}{2}\sin^2\theta\cos^2\theta + \frac{1}{2}\sin^4\theta + \sin^3\theta\cos\theta\right) \\ &= [\cos(2\theta) - \sin(2\theta)][1 + \sin(2\theta)]. \end{aligned} \tag{3.28}$$

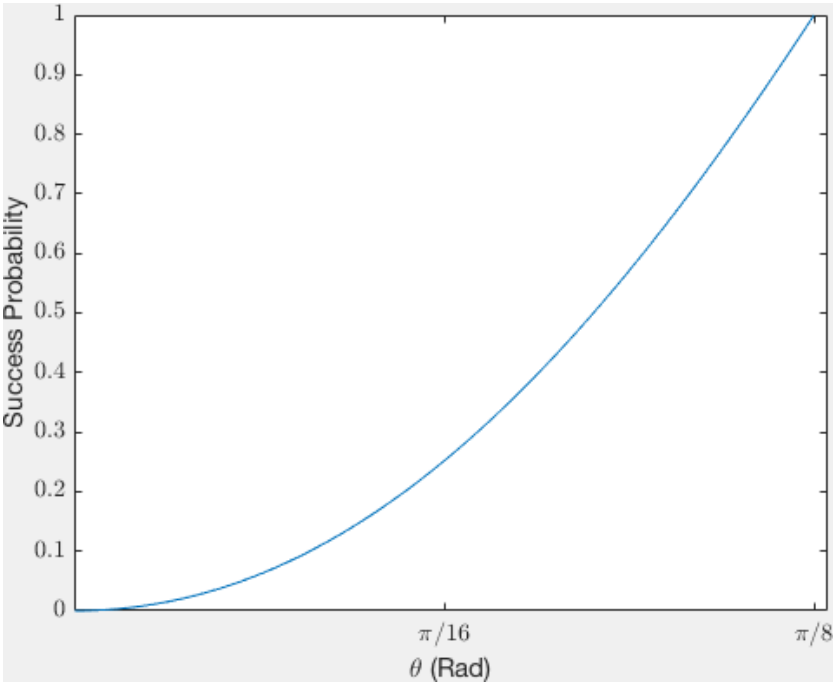


Figure 3.8: The success probability given by $1 - p_{f,1}$ where $p_{f,1}$ is given in (3.28). This just runs in the range from $0 \leq \theta \leq 22.5°(\pi/8$ rad$)$ as beyond that $c_{00} > 1/2$ and the measurement related to this success probability is not valid.

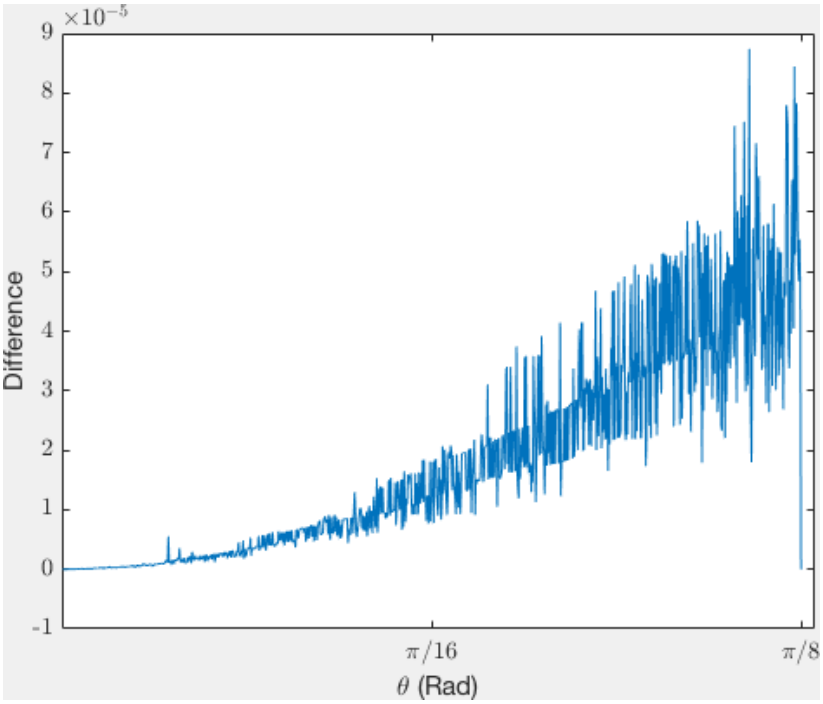


Figure 3.9: The difference between the success probability given by SDP and the group theory method. The maximum difference is about $5\text{x}10^{-5}$ *and the difference is larger for larger* θ*. The fluctuations are likely to be a resulting of numerical error.*

The formation of the measurement gives strong reason to believe it is optimal due to the optimisation of the coefficients c_{ij}. This combined with the very small gap between the bound from SDP and the success probability given in the group approach as can be seen in figure 3.9 further concretes the belief that there is not a more optimal method.

3.2 Eliminating One Of Four States When $\theta \geq 22.5°$

A recurring problem is that our methods eliminate with certainty when $\theta = 22.5°$ yet not for angles greater than this. PBR showed this is possible as they give a proof showing that for an angle θ the required number of qubits n for certain elimination must satisfy the inequality

$$2^{1/n} - 1 \leq \tan(\theta/2). \tag{3.29}$$

Re-arranging for θ we have

$$\theta \geq 2\tan^{-1}(2^{1/n} - 1). \tag{3.30}$$

For $n = 2$ we see that $\theta \geq 22.5°$. As this is an inequality for any value of $\theta \geq 22.5°$ we should be able to attain certain elimination.
One method to do this is using an ancillary qubit to dump the 'extra information' and

reduce the system down to the original PBR measurement with $\theta = 22.5°$.
First we couple the system qubits to the ancilla so that

$$|\theta\rangle|0\rangle \rightarrow |0\rangle|a\rangle \tag{3.31}$$

$$|-\theta\rangle|0\rangle \rightarrow |+\rangle|b\rangle \tag{3.32}$$

where

$$\langle a|b\rangle = \sqrt{2}\cos(2\theta). \tag{3.33}$$

The generalised states are then transformed as such,

$$|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle \rightarrow |00\rangle|aa\rangle, |0+\rangle|ab\rangle, |+0\rangle|ba\rangle, |++\rangle|bb\rangle, \tag{3.34}$$

and then we measure the system qubits using the PBR basis ignoring the ancilla qubits.

3.3 Unambiguously Eliminating Two Out Of Four States In The Two-Qubit Case

Eliminating two out of four states is different to the one of four case as now you require a basis that has components orthogonal to multiple different states. There are $\binom{4}{2} = 6$ different ways to eliminate two out of four states.

$$\begin{aligned}
A =& \{|\theta,\theta\rangle, |\theta,-\theta\rangle\},\\
B =& \{|\theta,\theta\rangle, |-\theta,\theta\rangle\},\\
C =& \{|\theta,-\theta\rangle, |-\theta,-\theta\rangle\},\\
D =& \{|-\theta,\theta\rangle, |-\theta,-\theta\rangle\},\\
E =& \{|\theta,-\theta\rangle, |-\theta,\theta\rangle\},\\
F =& \{|\theta,\theta\rangle, |-\theta,-\theta\rangle\}.
\end{aligned} \tag{3.35}$$

Before showing the derivation of the bound for the success probability to eliminate two out of four states I shall quickly talk about an interesting and slightly similar problem called the *Mean King Problem*.

Mean King Problem and Two Qubit States

This is a thought experiment proposed by Vaidman et al. [40] that was later reformed into a problem stated by Englert et al. [41] as:

A mean king challenges a physicist, who got stranded on the remote island ruled by the king to prepare a spin $1/2$ atom in any state of her choosing and to perform a control measurement of her liking. Between her measurement, the king's men determine the value of either σ_x, σ_y or σ_z. Only after she completed her control measurement, the physicist is told which spin component has been measured, and she must then state the result of that intermediate measurement correctly. How does she do it?

The solution with success probability one is given in the original paper [40] and involves preparing a composite system of two spin $1/2$ particles and measuring in an entangled basis. Then a photon analogue to the problem was given by [42] which measures a two qubit system not too dissimilar from our one to ascertain the values from the three mutually complimentary measurements. This is then formed into a QKD protocol [43]. There are similarities with this problem and our elimination problem in the attempt to obtain information from a two qubit system with six possible outcome states. Even though the desired outcome is different it is an interesting problem to look at.

Now I will present an analytic method of finding a bound proposed by my supervisor Erika Andersson [39] on the measurement that agrees with the bound from SDP and then give the measurement that saturates the bound and is hence the optimal measurement.

3.3.1 Bound

In the first four pairs $(A - D)$ from equation (3.35) either the first or the second qubit in each of the pairs is the same. For example in C both the second qubits are $|-\theta\rangle$. For outcome E we learn that the two qubits have a different state and for outcome F we learn they have the same state. Let us assume the state sent was $|\theta, \theta\rangle$ but all we know is that it is either $|\theta, \theta\rangle$ or $|-\theta, -\theta\rangle$, which occur with the same prior probability. The possible outcomes are anything that doesn't include $|\theta, \theta\rangle$, which is C, D and E. Outcomes C and D unambiguously distinguish between $|\theta, \theta\rangle$ and $|-\theta, -\theta\rangle$. This is because they both contain $|-\theta, -\theta\rangle$. The IDP limit [11][12][13] for unambiguous discrimination states that the probability of unambiguous discrimination between $|\theta, \theta\rangle$ and $|-\theta, -\theta\rangle$ can't exceed $1 - \langle -\theta, -\theta|\theta, \theta\rangle = 1 - \cos^2(2\theta)$. Therefore $p(C|\theta, \theta) + p(D|\theta, \theta) \leq 1 - \cos^2(2\theta)$.
In a similar vein say the state again was $|\theta, \theta\rangle$ but in this case we know it was either $|\theta, \theta\rangle$ or $|\theta, -\theta\rangle$. Then outcomes C and E unambiguously distinguish between $|\theta, \theta\rangle$ and $|\theta, -\theta\rangle$. This probability can't exceed $1 - \langle \theta, -\theta|\theta, \theta\rangle = 1 - \cos(2\theta)$ giving $p(C|\theta, \theta) + p(E|\theta, \theta) \leq 1 - \cos(2\theta)$.

Finally if we know the state is either $|\theta, \theta\rangle$ or $|-\theta, \theta\rangle$ then outcomes D and E unambiguously discriminate between them. Leaving us with the inequality

$$p(D|\theta,\theta) + p(E|\theta,\theta) \leq 1 - \cos(2\theta).$$

Now we have three inequalities,

$$\begin{aligned} p(C|\theta,\theta) + p(D|\theta,\theta) &\leq 1 - \cos^2(2\theta), \\ p(C|\theta,\theta) + p(E|\theta,\theta) &\leq 1 - \cos(2\theta), \\ p(D|\theta,\theta) + p(E|\theta,\theta) &\leq 1 - \cos(2\theta). \end{aligned} \tag{3.36}$$

Adding these up we obtain

$$2[p(C|\theta,\theta) + p(D|\theta,\theta) + p(E|\theta,\theta)] \leq 3 - 2\cos(2\theta) - \cos^2(2\theta) = 4 - [1 + \cos(2\theta)]^2. \tag{3.37}$$

If we require a result with 100% success probability then the failure probability will be zero and $p(C|\theta,\theta)+p(D|\theta,\theta)+p(E|\theta,\theta) = 1$. This then makes the LHS of the inequality in (3.37) two. Re-arranging for $\cos(2\theta)$ we get,

$$\begin{aligned} \cos(2\theta) &\leq \sqrt{2} - 1 \\ 2\theta &\leq 65.53^\circ, \end{aligned} \tag{3.38}$$

which agrees with the bound from SDP.

3.3.2 Sequential Approach

An approach Sarah Croke came up with [44], was to extend the ancilla approach used to eliminate one of four for $\theta \geq 22.5^\circ$ as shown in chapter (3.2). Once we eliminate one of four states we are left with three out of the four states $|aa\rangle, |ab\rangle, |ba\rangle, |bb\rangle$, depending on which state was eliminated in the first stage of the measurement. If we say for example $|-\theta,-\theta\rangle$ was eliminated, which means we have also eliminated $|bb\rangle$ as can be seen from (3.34), then if we can eliminate one of $|aa\rangle, |ab\rangle, |ba\rangle$ we have successfully eliminated two of the four states.

Sarah proposed the following special case,

$$\begin{aligned} |a\rangle &= |0\rangle, \\ |b\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle. \end{aligned} \tag{3.39}$$

In this case $\langle a|b\rangle = 1/\sqrt{3}$ and from equation (3.33) we see that $\cos(2\theta) = 1/\sqrt{6}$ or $\theta \simeq 37.95^\circ$. If $|-\theta,-\theta\rangle$ was eliminated we now are trying to eliminate one of

$$\begin{aligned} |aa\rangle &= |00\rangle, \\ |ab\rangle &= \frac{1}{\sqrt{3}}|00\rangle + \sqrt{\frac{2}{3}}|01\rangle, \\ |ba\rangle &= \frac{1}{\sqrt{3}}|00\rangle + \sqrt{\frac{2}{3}}|10\rangle. \end{aligned} \tag{3.40}$$

The following three states are orthogonal to $|aa\rangle, |ab\rangle, |ba\rangle$ respectively,

$$\begin{aligned}
|\overline{aa}\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \\
|\overline{ab}\rangle &= \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle, \\
|\overline{ba}\rangle &= \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{2}|01\rangle - \frac{1}{2}|10\rangle.
\end{aligned} \tag{3.41}$$

$|\overline{aa}\rangle, |\overline{ab}\rangle$ and $|\overline{ba}\rangle$ are orthonormal to each other and also all orthonormal to $|11\rangle$ which completes the basis. The state $|11\rangle$ never occurs though. This is because $|11\rangle$ is orthogonal to $|aa\rangle, |ab\rangle$ and $|ba\rangle$. Therefore we have guaranteed elimination 100% of the time. This unfortunately falls agonisingly short of the bound from equation (3.38) and SDP, which lies at $\cos(2\theta) = \sqrt{2} - 1$ or $\theta \approx 37.75°$.
Using the same sequential approach as above I labelled the states $|a\rangle$ and $|b\rangle$ with a separation angle from the bound in (3.38), which was equivalent to that from SDP. In a general case we have

$$|a\rangle = |0\rangle \qquad \text{and} \qquad |b\rangle = b_1|0\rangle + b_2|1\rangle, \tag{3.42}$$

and therefore

$$\begin{aligned}
|aa\rangle &= |00\rangle, \\
|ab\rangle &= b_1|00\rangle + b_2|01\rangle, \\
|ba\rangle &= b_1|00\rangle + b_2|10\rangle.
\end{aligned} \tag{3.43}$$

Then by requiring orthogonality we get the orthogonal states as

$$\begin{aligned}
|\overline{aa}\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \\
|\overline{ab}\rangle &= \sqrt{\frac{b_1^2}{b_2^2 + 2b_1^2}}(|01\rangle - |10\rangle + \frac{b_2}{b_1}|00\rangle), \\
|\overline{ba}\rangle &= \sqrt{\frac{b_1^2}{b_2^2 + 2b_1^2}}(|10\rangle - |01\rangle + \frac{b_2}{b_1}|00\rangle).
\end{aligned} \tag{3.44}$$

If we wish to test this approach with the bound from (3.38), which was equivalent to that from SDP then we require the overlap

$$\langle a|b\rangle = \sqrt{2}\cos(2\theta) = \sqrt{2}(\sqrt{2} - 1) = 2 - \sqrt{2}, \tag{3.45}$$

therefore

$$b_1 = 2 - \sqrt{2} \qquad \text{and} \qquad b_2 = \sqrt{1 - b_1^2} = \sqrt{5 - 4\sqrt{2}}. \tag{3.46}$$

The coefficients for each term in each situation are then given as

General	b_1	b_2
Sarah's	$1/\sqrt{3}$	$\sqrt{2/3}$
Bound from (3.38)	$2-\sqrt{2}$	$\sqrt{5-4\sqrt{2}}$.

Once we sum the operators (not including the measurement operator corresponding to the inconclusive result and discarding the space given by $|11\rangle$) onto the orthogonal states we obtain,

$$|\overline{aa}\rangle\langle\overline{aa}| + |\overline{ab}\rangle\langle\overline{ab}| + |\overline{ba}\rangle\langle\overline{ba}| = \begin{pmatrix} \frac{2b_2^2}{b_1^2((b_2/b_1)^2+2)} & 0 & 0 \\ 0 & \frac{2}{b_1^2((b_2/b_1)^2+2)} + \frac{1}{2} & \frac{1}{2} - \frac{2}{b_1^2((b_2/b_1)^2+2)} \\ 0 & \frac{1}{2} - \frac{2}{b_1^2((b_2/b_1)^2+2)} & \frac{2}{b_1^2((b_2/b_1)^2+2)} + \frac{1}{2} \end{pmatrix}. \tag{3.47}$$

If we substitute in Sarah's coefficients then we obtain the identity as expected and therefore have a guaranteed elimination, whereas if we substitute in our bound then we get

$$|\overline{aa}\rangle\langle\overline{aa}| + |\overline{ab}\rangle\langle\overline{ab}| + |\overline{ba}\rangle\langle\overline{ba}| = \begin{pmatrix} 0.978 & 0 & 0 \\ 0 & 1.011 & -0.011 \\ 0 & -0.011 & 1.011 \end{pmatrix}. \tag{3.48}$$

This does not give us the identity matrix as required for a certain measurement, even if we scale it and so a measurement of this form will not successfully eliminate two out of four states 100% of the time.

Caves, Fuchs and Schack [8] also found limitations on the ability of eliminating one state out of three pure states. The authors look to find under what conditions is there a measurement whose outcome contradicts one of the three distinct, non-orthogonal pure states. This is equivalent to our elimination requirement in that we need a measurement that rules one of the possible states.

For three pure states $|\psi_1\rangle, |\psi_2\rangle$ and $|\psi_3\rangle$, the overlaps are defined as,

$$\begin{aligned} a &\equiv |\langle\psi_1|\psi_2\rangle|^2, \\ b &\equiv |\langle\psi_2|\psi_3\rangle|^2, \\ c &\equiv |\langle\psi_3|\psi_1\rangle|^2. \end{aligned} \tag{3.49}$$

The conclusion to this section of the authors work is that if the conditions

$$\begin{aligned} &a+b+c \leq 1, \\ &(a+b+c-1)^2 \geq 4abc, \end{aligned} \tag{3.50}$$

are met then a measurement exists that will eliminate one of the three states unambiguously with certainty. For Sarah's method we have,

$$\begin{aligned}
|\psi_1\rangle &= |aa\rangle = |00\rangle, \\
|\psi_2\rangle &= |ab\rangle = \frac{1}{\sqrt{3}}|00\rangle + \sqrt{\frac{2}{3}}|01\rangle, \\
|\psi_3\rangle &= |ba\rangle = \frac{1}{\sqrt{3}}|00\rangle + \sqrt{\frac{2}{3}}|10\rangle.
\end{aligned} \tag{3.51}$$

This gives us

$$a = \frac{1}{3}, \qquad b = \frac{1}{9} \qquad \text{and} \qquad c = \frac{1}{3}, \tag{3.52}$$

from which the conditions are

$$\begin{aligned}
&a + b + c = \frac{7}{9} < 1, \\
&(a + b + c - 1)^2 = \frac{4}{81} \qquad \text{and} \qquad 4abc = \frac{4}{81}.
\end{aligned} \tag{3.53}$$

So the bottom bound in (3.50) is saturated when $\theta \simeq 37.95$ whilst the top bound is satisfied. Using the bound in (3.38) we obtained from SDP, we can see if it is possible to eliminate one of the three remaining states by checking against the conditions in [8]. The states are now

$$\begin{aligned}
|\psi_1\rangle &= |aa\rangle = |00\rangle, \\
|\psi_2\rangle &= |ab\rangle = 2 - \sqrt{2}|00\rangle + \sqrt{5 - 4\sqrt{2}}|01\rangle, \\
|\psi_3\rangle &= |ba\rangle = 2 - \sqrt{2}|00\rangle + \sqrt{5 - 4\sqrt{2}}|10\rangle,
\end{aligned} \tag{3.54}$$

giving us,

$$a = 6 - 4\sqrt{2}, \qquad b = 68 - 48\sqrt{2} \qquad \text{and} \qquad c = 6 - 4\sqrt{2}. \tag{3.55}$$

The conditions are then,

$$\begin{aligned}
&a + b + c \simeq 0.80 < 1, \\
&(a + b + c - 1)^2 \simeq 0.038 \qquad \text{and} \qquad 4abc \simeq 0.055.
\end{aligned} \tag{3.56}$$

The second condition from (3.50) is not met and so it is not possible to eliminate one of the states in (3.54) with certainty. This does not mean that it is impossible to perform a two out of four elimination for $\theta \simeq 37.75°$. It just means that we can't do the initial ancilla method as this must remove some information. It also doesn't mean that a sequential measurement is impossible, but we are yet to find a method to reach the bound eliminating the states one qubit at a time.

3.3.3 Optimal Measurement

In this section we will derive the measurement that was found to eliminate two of the four states 100% of the time.

If we look at obtaining outcomes E and F from (3.35) then an orthogonal (but unnormalised) basis for the space spanned by $|\theta,\theta\rangle$ and $|-\theta,-\theta\rangle$ can be given by

$$\{\cos^2\theta|00\rangle + \sin^2\theta|11\rangle,\ \ |01\rangle + |10\rangle\}. \tag{3.57}$$

This is due to $|\theta,\theta\rangle$ and $|-\theta,-\theta\rangle$ being,

$$\cos^2\theta|00\rangle + \sin^2\theta|11\rangle \pm \cos\theta\sin\theta(|01\rangle + |10\rangle). \tag{3.58}$$

Similarly, an orthogonal (but unnormalised) basis for the space spanned by the states $|\theta,-\theta\rangle$ and $|-\theta,\theta\rangle$ is

$$\{\cos^2\theta|00\rangle - \sin^2\theta|11\rangle,\ \ |01\rangle - |10\rangle\}. \tag{3.59}$$

This is due to $|\theta,\theta\rangle$ and $|-\theta,-\theta\rangle$ being,

$$\cos^2\theta|00\rangle - \sin^2\theta|11\rangle \pm \cos\theta\sin\theta(|01\rangle - |10\rangle). \tag{3.60}$$

To have an unambiguous measurement we need to form measurement operators that are proportional to projectors onto $|01\rangle - |10\rangle$ and $\sin^2\theta|00\rangle - \cos^2\theta|11\rangle$ to eliminate F. To eliminate E we we need to form measurement operators that are proportional to projectors onto $|01\rangle + |10\rangle$ and $\sin^2\theta|00\rangle + \cos^2\theta|11\rangle$.

If projectors onto $|01\rangle \pm |10\rangle$ have the same weight α then the contribution to the sum of all measurement operators will be

$$2\alpha(|01\rangle\langle 01| + |10\rangle\langle 10|). \tag{3.61}$$

If projectors onto $\cos^2\theta|00\rangle \pm \sin^2\theta|11\rangle$ have the same weight β then there contribution to the sum is

$$2\beta(\sin^4\theta|00\rangle\langle 00| + \cos^4\theta|11\rangle\langle 11|). \tag{3.62}$$

For $0 \leq \theta \leq 45^\circ$ it holds that $\sin^4\theta \leq \cos^4\theta$. For a measurement with perfect success we require the other operators to have a greater contribution to $|00\rangle\langle 00|$ than to $|11\rangle\langle 11|$. This is needed so the sum can reach the identity which is required for the measurement to eliminate one the pairs every single time. We shall label these spaces with

$$|\psi^{\pm}_{cs}\rangle = \sin^2|00\rangle \pm \cos^2\theta|11\rangle \qquad \text{and} \qquad |\psi^{\pm}_{01}\rangle = |01\rangle \pm |10\rangle. \tag{3.63}$$

The pair A spans the space $|\theta\rangle\langle\theta| \otimes I_2$, where I_2 is simply the identity operator in the space of the second qubit. Therefore we require

$$\Pi_{\overline{A}} \propto |\overline{\theta}\rangle_{11}\langle\overline{\theta}| \otimes \pi_2, \tag{3.64}$$

where π_2 has to be some operator acting on the second qubit. It makes sense for $\pi_2 \propto |0\rangle_{22}\langle 0|$ to make up for some of the shortfall in $|00\rangle\langle 00|$.
The pair B spans the pace $I_1 \otimes |\theta\rangle\langle\theta|$ so then we require

$$\Pi_{\overline{B}} \propto |\overline{\theta}\rangle_{22}\langle\overline{\theta}| \otimes \pi_1, \tag{3.65}$$

again choosing $\pi_1 \propto |0\rangle_{11}\langle 0|$. The measurement operators can then be given by

$$\Pi_{\overline{A}} = \gamma|\overline{\theta}, 0\rangle\langle\overline{\theta}, 0|, \quad \Pi_{\overline{B}} = \gamma|0, \overline{\theta}\rangle\langle 0, \overline{\theta}|, \tag{3.66}$$

where γ is the relevant weighting For the final two operators we have,

$$\begin{aligned} \Pi_{\overline{C}} &\propto |0, \overline{-\theta}\rangle\langle 0, \overline{-\theta}|, \\ \Pi_{\overline{D}} &\propto |\overline{-\theta}, 0\rangle\langle\overline{-\theta}, 0|, \end{aligned} \tag{3.67}$$

where again we have used a projection onto $|0\rangle\langle 0|$ for either the first or second qubit as we did for $\Pi_{\overline{A}}$ and $\Pi_{\overline{B}}$. From the fact

$$|\overline{\theta}\rangle\langle\overline{\theta}| + |\overline{-\theta}\rangle\langle\overline{-\theta}| = 2\sin^2\theta|0\rangle\langle 0| + 2\cos^2\theta|1\rangle\langle 1|, \tag{3.68}$$

we can see the contribution of the measurement operators $\Pi_{\overline{A-D}}$ with there weighting γ, to the sum of measurement operators is given by

$$\gamma[4\sin^2\theta|00\rangle\langle 00| + 2\cos^2\theta(|01\rangle\langle 01| + |10\rangle\langle 10|)]. \tag{3.69}$$

If we sum all the measurement operators up then we get

$$\begin{aligned} &(2\beta\sin^4\theta + 4\gamma\sin^2\theta)|00\rangle\langle 00| + (2\alpha + 2\gamma\cos^2\theta)|01\rangle\langle 01| + \\ &(2\alpha + 2\gamma\cos^2\theta)|10\rangle\langle 10| + 2\beta\cos^4\theta|11\rangle\langle 11|. \end{aligned} \tag{3.70}$$

For it to be a valid measurement we require:

$$\begin{aligned} 2\beta\sin^4\theta + 4\gamma\sin^2\theta &\leq 1, \\ 2\beta\cos^4\theta &\leq 1, \\ 2\alpha + 2\gamma\cos^2\theta &\leq 1. \end{aligned} \tag{3.71}$$

When all the inequalities are saturated simultaneously then the sum of the measurement operators becomes the identity and our measurement will have guaranteed success. A method to saturate the bounds is by setting

$$\beta = \frac{1}{2\cos^4\theta}, \tag{3.72}$$

to saturate the second bound from (3.71) then looking at the first bound we have

$$\begin{aligned} \tan^4\theta + 4\gamma\sin^2\theta &= 1, \\ \gamma &= \frac{1 - \tan^4\theta}{4\sin^2\theta}. \end{aligned} \tag{3.73}$$

To saturate the third bound from equation (3.71) we need to solve for α in terms of the β and γ values we have used above.

$$2\alpha + 2\cos^2\theta\left(\frac{1-\tan^4\theta}{4\sin^2\theta}\right) = 2\alpha + \frac{1}{2}(\cot^2\theta - \tan^2\theta) = 1,$$
$$\alpha = \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) = \frac{1}{2} - \gamma\cos^2\theta. \tag{3.74}$$

The measurement operators for the six outcomes A to F then become

$$\begin{aligned}
\Pi_{\overline{A}} =& \gamma|\overline{\theta},0\rangle\langle\overline{\theta},0|, \quad \Pi_{\overline{B}} = \gamma|0,\overline{\theta}\rangle\langle 0,\overline{\theta}|,\\
\Pi_{\overline{C}} =& \gamma|0,\overline{-\theta}\rangle\langle 0,\overline{-\theta}|, \quad \Pi_{\overline{D}} = \gamma|\overline{-\theta},0\rangle\langle\overline{-\theta},0|,\\
\Pi_{\overline{E}} =& \alpha|\psi_{01}^{+}\rangle\langle\psi_{01}^{+}| + \beta|\psi_{cs}^{+}\rangle\langle\psi_{cs}^{+}|\\
\Pi_{\overline{F}} =& \alpha|\psi_{01}^{-}\rangle\langle\psi_{01}^{-}| + \beta|\psi_{cs}^{-}\rangle\langle\psi_{cs}^{-}|.
\end{aligned} \tag{3.75}$$

If we define the column vectors as

$$|00\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ |01\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ |10\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ |11\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \tag{3.76}$$

we can represent these operators in matrix form using only the variable θ. For $\Pi_{\overline{A-F}}$ we have,

$$\Pi_{\overline{A}} = \begin{pmatrix} \frac{1}{4}(1-\tan^4\theta) & 0 & -\frac{1}{4}(\cot\theta - \tan^3\theta) & 0\\ 0 & 0 & 0 & 0\\ -\frac{1}{4}(\cot\theta - \tan^3\theta) & 0 & \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0\\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.77}$$

$$\Pi_{\overline{B}} = \begin{pmatrix} \frac{1}{4}(1-\tan^4\theta) & -\frac{1}{4}(\cot\theta - \tan^3\theta) & 0 & 0\\ -\frac{1}{4}(\cot\theta - \tan^3\theta) & \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.78}$$

$$\Pi_{\overline{C}} = \begin{pmatrix} \frac{1}{4}(1-\tan^4\theta) & \frac{1}{4}(\cot\theta - \tan^3\theta) & 0 & 0\\ \frac{1}{4}(\cot\theta - \tan^3\theta) & \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.79}$$

$$\Pi_{\overline{D}} = \begin{pmatrix} \frac{1}{4}(1-\tan^4\theta) & 0 & \frac{1}{4}(\cot\theta - \tan^3\theta) & 0\\ 0 & 0 & 0 & 0\\ \frac{1}{4}(\cot\theta - \tan^3\theta) & 0 & \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0\\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.80}$$

$$\Pi_{\overline{E}} = \begin{pmatrix} \frac{1}{2}\tan^4\theta & 0 & 0 & \frac{1}{2}\tan^2\theta \\ 0 & \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) & \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0 \\ 0 & \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) & \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0 \\ \frac{1}{2}\tan^2\theta & 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad (3.81)$$

$$\Pi_{\overline{F}} = \begin{pmatrix} \frac{1}{2}\tan^4\theta & 0 & 0 & -\frac{1}{2}\tan^2\theta \\ 0 & \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) & -\frac{1}{2} + \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0 \\ 0 & -\frac{1}{2} + \frac{1}{4}(\cot^2\theta - \tan^2\theta) & \frac{1}{2} - \frac{1}{4}(\cot^2\theta - \tan^2\theta) & 0 \\ -\frac{1}{2}\tan^2\theta & 0 & 0 & \frac{1}{2} \end{pmatrix}. \quad (3.82)$$

If we sum up all these operators we obtain the identity. All the values for α, β and γ satisfy the inequalities in (3.71) as long as $\cos(2\theta) \leq \sqrt{2} - 1$. This gives us a measurement that succeeds 100% of the time at eliminating two out of four states for when $\theta \geq 32.76°$. With the measurement operators we can calculate the overall probability using

$$p(i|\rho) = Tr(\Pi_i \rho), \quad (3.83)$$

where ρ is the density matrix of the prepared state and is given by

$$\rho = \sum_i p_i \rho_i = \begin{pmatrix} \cos^4\theta & 0 & 0 & 0 \\ 0 & \cos^2\theta\sin^2\theta & 0 & 0 \\ 0 & 0 & \cos^2\theta\sin^2\theta & 0 \\ 0 & 0 & 0 & \sin^4\theta \end{pmatrix}, \quad (3.84)$$

where p_i are the prior probabilities of the states described by ρ_i. In our case we have an equal probability of each initial state so $p_i = 1/4$. The calculated probabilities are

$$\begin{aligned} p(A|\rho) = p(B|\rho) = p(C|\rho) = p(D|\rho) &= \frac{1}{2}\cos(2\theta), \\ p(E|\rho) = p(F|\rho) &= \frac{1}{2} - \cos(2\theta). \end{aligned} \quad (3.85)$$

Summing up the probabilities we get

$$4(\frac{1}{2}\cos(2\theta)) + 2(\frac{1}{2} - \cos(2\theta)) = 1, \quad (3.86)$$

as expected. The results of this measurement only hold for $\cos(2\theta) \leq \sqrt{2} - 1$. It is also worthwhile introducing the probability of each outcome given a certain initial preparation and these are shown in table 3.1. In the table we see that each outcome gives a probability of 0 for two of the preparations as these are the states that outcome eliminates. If we look at the the initial state being prepared in $|\theta, \theta\rangle$ then we see A, B and F have zero probability of being the outcome, which is required as they all eliminate $|\theta, \theta\rangle$. Then C and D have a probability of $\cos(2\theta)$. These two outcomes eliminate $|\theta, -\theta\rangle, |-\theta, -\theta\rangle$

and $|-\theta,\theta\rangle, |-\theta,-\theta\rangle$ respectively, whereas E eliminates $|\theta,-\theta\rangle, |-\theta,\theta\rangle$. We can see that in the two combinations C and D eliminate only one qubit is the same as the prepared state (there is only one $|\theta\rangle$) and they both eliminate the state $|-\theta,-\theta\rangle$. As the prepared state is $|\theta,\theta\rangle$ the most likely state to be eliminated is $|-\theta,-\theta\rangle$ and this then explains why outcome E is less likely than outcomes C or D. The same logic can be applied to the other initial states and explains the symmetry about A, B, C, D and then separately E and F. We also see these results are comparable to those from equation (3.85).

	$\|\theta,\theta\rangle$	$\|\theta,-\theta\rangle$	$\|-\theta,\theta\rangle$	$\|-\theta,-\theta\rangle$
A	0	0	$\cos(2\theta)$	$\cos(2\theta)$
B	0	$\cos(2\theta)$	0	$\cos(2\theta)$
C	$\cos(2\theta)$	0	$\cos(2\theta)$	0
D	$\cos(2\theta)$	$\cos(2\theta)$	0	0
E	$1-2\cos(2\theta)$	0	0	$1-2\cos(2\theta)$
F	0	$1-2\cos(2\theta)$	$1-2\cos(2\theta)$	0

Table 3.1: A table showing the probabilities of each measurement outcome given the initial preparation state. The left column has the different outcomes and the top row showing the different initial states. So the top left box represents $p(A|\theta,-\theta\rangle) = 0$.

We can also write the measurement in terms of the pure states the measurement operators project upon.

$$\begin{aligned}
|\psi_{\overline{A}}\rangle &= \sqrt{\gamma}\sin\theta|00\rangle - \sqrt{\gamma}\cos\theta|10\rangle,\\
|\psi_{\overline{B}}\rangle &= \sqrt{\gamma}\sin\theta|00\rangle - \sqrt{\gamma}\cos\theta|01\rangle,\\
|\psi_{\overline{C}}\rangle &= \sqrt{\gamma}\sin\theta|00\rangle + \sqrt{\gamma}\cos\theta|01\rangle,\\
|\psi_{\overline{D}}\rangle &= \sqrt{\gamma}\sin\theta|00\rangle + \sqrt{\gamma}\cos\theta|10\rangle,\\
|\psi_{\overline{E}}\rangle &= \sqrt{\beta}\sin^2\theta|00\rangle + \sqrt{\alpha}|01\rangle + \sqrt{\alpha}|10\rangle + \sqrt{\beta}\cos^2\theta|11\rangle,\\
|\psi_{\overline{F}}\rangle &= \sqrt{\beta}\sin^2\theta|00\rangle + \sqrt{\alpha}|01\rangle - \sqrt{\alpha}|10\rangle - \sqrt{\beta}\cos^2\theta|11\rangle,
\end{aligned} \tag{3.87}$$

and in matrix form this is

$$\begin{pmatrix}
\sqrt{\gamma}\sin\theta & 0 & -\sqrt{\gamma}\cos\theta & 0\\
\sqrt{\gamma}\sin\theta & -\sqrt{\gamma}\cos\theta & 0 & 0\\
\sqrt{\gamma}\sin\theta & \sqrt{\gamma}\cos\theta & 0 & 0\\
\sqrt{\gamma}\sin\theta & 0 & \sqrt{\gamma}\cos\theta & 0\\
\sqrt{\beta}\sin^2\theta & \sqrt{\alpha} & \sqrt{\alpha} & \sqrt{\beta}\cos^2\theta\\
\sqrt{\beta}\sin^2\theta & \sqrt{\alpha} & -\sqrt{\alpha} & -\sqrt{\beta}\cos^2\theta
\end{pmatrix}. \tag{3.88}$$

Now considering the largest angle for two state exclusion to be allowed ($\cos(2\theta) = \sqrt{2} - 1 \rightarrow \theta = 32.76^\circ$ and using standard trigonometric identities we have

$$\begin{aligned}
&\cos(2\theta) = \sqrt{2} - 1 = 2\cos^2\theta - 1, \\
&\cos^2\theta = \frac{1}{\sqrt{2}}, \\
&\sin^2\theta = 1 - \frac{1}{\sqrt{2}}, \\
&\tan^2\theta = \frac{\sin^2\theta}{\cos^2\theta} = \frac{1 - 1/\sqrt{2}}{1/\sqrt{2}} = \sqrt{2} - 1.
\end{aligned} \tag{3.89}$$

From these we can calculate the coefficients α, β and γ as

$$\begin{aligned}
&\beta = \frac{1}{2\cos^4\theta} = \frac{1}{2.1/2} = 1, \\
&\gamma = \frac{1 - \tan^4\theta}{4\sin^2\theta} = \frac{1 - (\sqrt{2}-1)^2}{4(1 - 1/\sqrt{2})} = \frac{1 - (3 - 2\sqrt{2})}{4(\sqrt{2}-1)/\sqrt{2}} = \frac{2\sqrt{2}(\sqrt{2}-1)}{4(\sqrt{2}-1)} = \frac{1}{\sqrt{2}}, \\
&\alpha = \frac{1}{2} - \gamma\cos^2\theta = \frac{1}{2} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} = 0.
\end{aligned} \tag{3.90}$$

3.3.4 Measurement With A Failure Probability

In the case when $\cos(2\theta) > \sqrt{2} - 1$ the weightings,

$$\begin{aligned}
\beta &= \frac{1}{2\cos^4\theta}, \\
\gamma &= \frac{1}{2\cos^2\theta}, \\
\alpha &= 0,
\end{aligned} \tag{3.91}$$

are chosen to satisfy the inequalities in (3.71). This saturates the bottom two inequalities, whilst satisfying the top one. Now the measurement operators are given by,

$$\Pi_A = \frac{1}{2}\begin{pmatrix} \tan^2\theta & 0 & -\tan^2\theta & 0 \\ 0 & 0 & 0 & 0 \\ -\tan^2\theta & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \Pi_B = \frac{1}{2}\begin{pmatrix} \tan^2\theta & -\tan^2\theta & 0 & 0 \\ -\tan^2\theta & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Pi_C = \frac{1}{2}\begin{pmatrix} \tan^2\theta & \tan^2\theta & 0 & 0 \\ \tan^2\theta & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \Pi_D = \frac{1}{2}\begin{pmatrix} \tan^2\theta & 0 & \tan^2\theta & 0 \\ 0 & 0 & 0 & 0 \\ \tan^2\theta & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Pi_E = \frac{1}{2}\begin{pmatrix} \tan^4\theta & 0 & 0 & \tan^2\theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \tan^2\theta & 0 & 0 & 1 \end{pmatrix}, \qquad \Pi_F = \frac{1}{2}\begin{pmatrix} \tan^4\theta & 0 & 0 & -\tan^2\theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\tan^2\theta & 0 & 0 & 1 \end{pmatrix}. \tag{3.92}$$

The failure operator is then the identity minus the sum of $\Pi_{A..F}$,

$$\Pi_{fail} = \begin{pmatrix} 1 - 2\tan^2\theta - \tan^4\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.93}$$

As it is just non-zero in the $|00\rangle\langle 00|$ space the failure probability is given by,

$$p_{fail} = \cos^4\theta(1 - 2\tan^2\theta - \tan^4\theta) = \cos^4\theta - 2\sin^2\theta\cos^2\theta - \sin^4\theta = 2\cos^4\theta - 1, \tag{3.94}$$

assuming the prior probabilities for states $|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta,-\theta\rangle$ are all $1/4$. The probabilities of each outcome are,

$$p(A) = p(B) = p(C) = p(D) = \sin^2\cos^2\theta, \tag{3.95}$$

$$p(E) = p(F) = \sin^4\theta. \tag{3.96}$$

The sum of these is

$$4p(A) + 2p(E) = 4\sin^2\theta\cos^2\theta + 2\sin^4\theta = 2 - 2\cos^4\theta = 1 - p_{fail} = p_s. \tag{3.97}$$

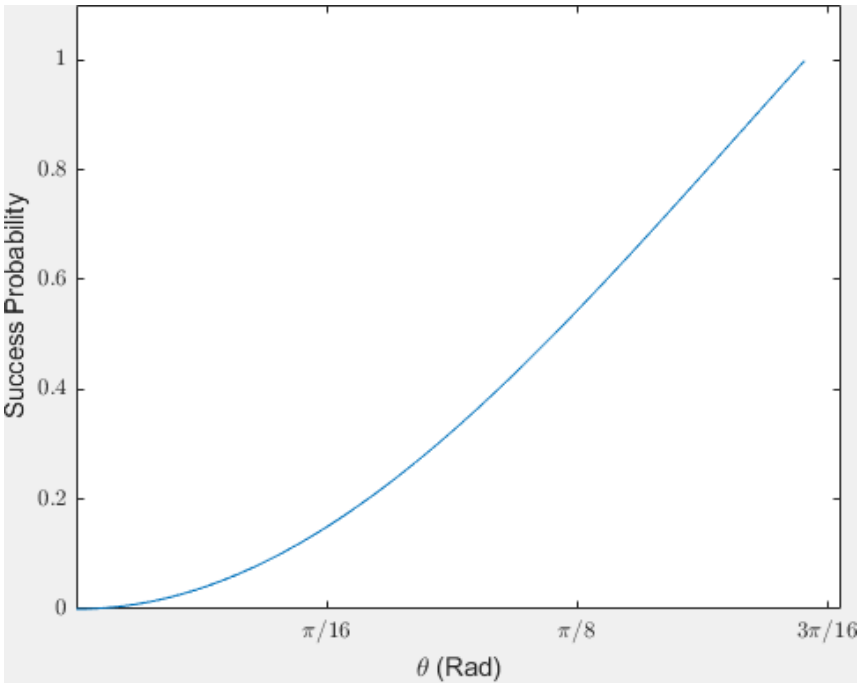


Figure 3.10: The success probability $p_s = 2 - 2\cos^4\theta$ of the two out of four elimination measurement described above for $\theta \leq 32.76°$.

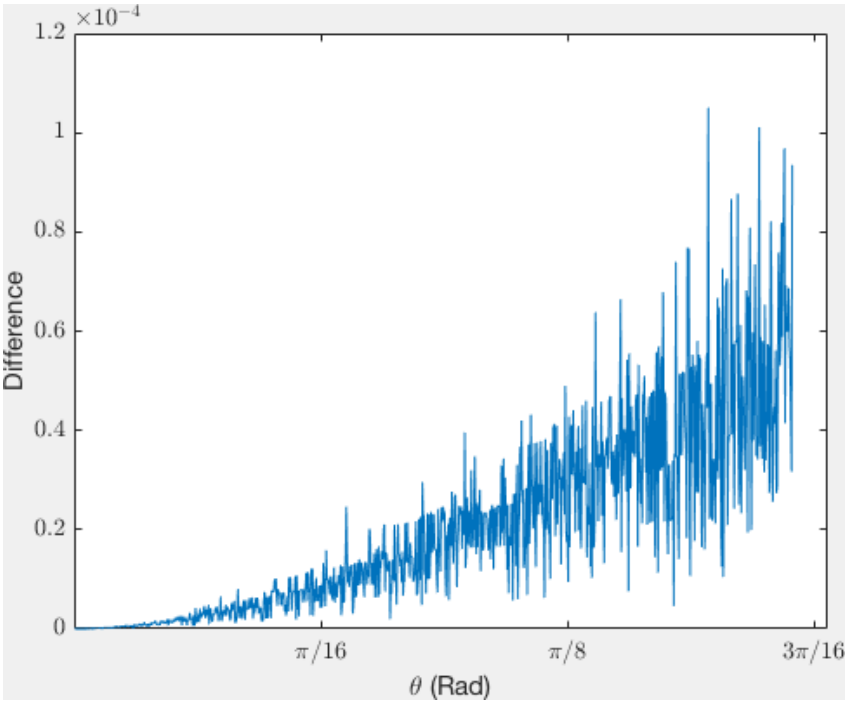


Figure 3.11: The difference between the success probabilities from the two out of four elimination measurement described above and the bound given by SDP, where the SDP bound is the larger value and y axis represents the difference. The spikes we believe are numerical error due to the randomness and the general trend is for the difference to increase with θ reaching a maximum of around $5x10^{-5}$.

The difference between the SDP bound and the measurement as shown in figure 3.11 is very small and the difference seems to be similar to what we expect due to the weak duality of the unambiguous SDP program. This leads me to believe the measurement is optimal for the unambiguous elimination of two of four states.

3.4 Application Of Elimination Measurement

In this section I will give a quick introduction to quantum cryptography including quantum key distribution and then proceed to explain how our measurement that eliminates two out the four states in the two-qubit case could potentially be used for this purpose.

3.4.1 Cryptography

One possible application is to use this measurement as the basis for a quantum key distribution (QKD) protocol. QKD is a method of distributing keys between participants, whom we shall call Alice and Bob. The key is a classical bit string that can be used to encrypt information so that it can be sent over an insecure channel, without the risk of it being read. Quantum mechanics is used to increase the security of the key and often with

the aim of making the protocols information-theoretically secure, which means that given unlimited computing power it is impossible to do better than a random guess. Currently some of our classical key distribution systems are assumed to be computationally hard to break, based on protocols such as RSA, elliptic curve and lattice-based cryptography [45][46][47]. Some of these are at risk with the growing potential of quantum computers, and QKD is a possible solution to that problem. The most well known QKD protocol is BB84 [48], but there are others QKD protocols that aim to make the protocols more experimentally feasible and also reduce the security risk involved in the practical aspect of the protocol [49]. Even though the BB84 key distribution method is information-theoretically secure, it still requires the practical apparatus to not leak information. Studies have found it is possible to break the encryption by gaining information from apparatus such as detectors [50]. This has lead to new protocols that are more resilient to attacks for example measurement device independent QKD [51] as well as better experimental equipment.

3.4.2 BB84

BB84 is named after its creators Bennett and Brassard and the year of its introduction in 1984. The fundamental principle that provides the security is the inability to distinguish between non-orthogonal quantum states and subsequently or equivalently the no-cloning theorem [52]. The aim is to produce a secret bit string shared between Alice and Bob. To do this Alice starts by sending one of the four following states on an insecure quantum channel (anybody can listen in/have access to the states),

$$|0\rangle,\ |1\rangle,\ |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\ \text{ and }\ |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{3.98}$$

where Bob chooses to measure in the $\{|0\rangle, |1\rangle\}$ or the $\{|+\rangle, |-\rangle\}$ basis. A measurement result by Bob of states $|0\rangle$ or $|-\rangle$ corresponds to the bit value 0 and a measurement of $|1\rangle$ or $|+\rangle$ corresponds to the bit value 1.
If he chooses the correct basis he will get the correct bit value, and if he measures in the wrong basis there is a 50% chance of getting the correct bit value as,

$$|\langle +|0\rangle|^2 = |\langle -|0\rangle|^2 = |\langle +|1\rangle|^2 = |\langle -|1\rangle|^2 = \frac{1}{2}. \tag{3.99}$$

Bob then announces publicly that he has performed the measurements and after this Alice will announce which basis the states were sent in. Bob can then tell Alice over an authenticated channel the positions of the bits he measured in the correct basis and then the results from those measurements are the shared secret bit string for the key. The results from when Bob measured in the wrong basis will be discarded.

To test the security we will introduce an eavesdropper named Eve. Eve's job is to try and obtain the secret key without Alice or Bob knowing. One option is to clone the states

sent to Bob and when Alice publicly declares the basis she measured in for each bit, Eve can measure afterwards and obtain the key. This is not possible deterministically due to the no-cloning theorem [52] that states that a single unknown quantum state can not be cloned with perfect success. Instead of cloning, Eve could measure the state then pass on the remaining state to Bob. This would be fine if Eve measured in the correct basis but if she chose the wrong basis she would project the state into the other basis and then send Bob a different state to that which Alice sent.

To check if the states have been tampered with Alice and Bob publicly compare a portion of the shared secret bit string and check whether the results match. Assuming that the implementation is perfect the only differences would be caused by interference from Eve. If differences occur and these are below a certain threshold Alice and Bob can perform error correction and privacy amplification to reduce Eve's information, otherwise the whole key is discarded and the process is started again. These are very simple cheating methods and many more complicated and better strategies exist, yet it can be proven that whatever strategy Eve attempts it is possible to bound how much information she has about the final secret shared key [53] [54].

3.4.3 B92

The B92 protocol [55] was introduces by Charles Bennett in 1992 and is a slightly simplified version of the BB84 protocol in that only two possible states are sent. This is either $|0\rangle$ or $|+\rangle$ where $|0\rangle$ would represent the bit 0 and $|+\rangle$ the bit 1. Bob then randomly picks the basis and if he can determine the bit from his measurement he keeps it, if the result is inconclusive he discards it. This is in essence an elimination measurement as a valid bit for example is when Bob measures in the diagonal basis and obtains the result relating to $|-\rangle$ then he knows it could not of been $|+\rangle$ sent and so it must have been $|0\rangle$.

Once the measurements are completed he tells Alice which bits were kept and this becomes the key. In contrast to BB84 the sifting is finished once Bob announces whether his results were conclusive and Alice is not required in this stage. Unfortunately despite being simpler this protocol seems to be less secure as it is sensitive to splitting attacks and the requirement for a strong reference pulse was even noted by the author in the initial paper. This leads to the B92 protocol having generally lower secure key rates than that of BB84 [49, 56].

3.4.4 Elimination QKD Protocol

Here we propose a method of using our two out of four elimination measurement as the basis for a QKD protocol. Alice sends one of two-qubit states

$|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta-,\theta\rangle$ to Bob. Bob performs the two out of four elimination measurement as described in section 3.3.3 to eliminate two of the possible states. Bob's possible outcomes are

$$\begin{aligned} A &= \{|\theta,\theta\rangle, |\theta,-\theta\rangle\}, \\ B &= \{|\theta,\theta\rangle, |-\theta,\theta\rangle\}, \\ C &= \{|\theta,-\theta\rangle, |-\theta,-\theta\rangle\}, \\ D &= \{|-\theta,\theta\rangle, |-\theta,-\theta\rangle\}, \\ E &= \{|\theta,-\theta\rangle, |-\theta,\theta\rangle\}, \\ F &= \{|\theta,\theta\rangle, |-\theta,-\theta\rangle\}, \end{aligned} \tag{3.100}$$

and for $\cos(2\theta) \geq \sqrt{2} - 1$ he will achieve one of the results $A - F$ every time. Each result will give a single bit of information that we can use for the shared secret key. For example if Bob obtains outcome A then he has eliminated both possibilities with $|\theta\rangle$ in the first qubit position. This means he has learnt that the first qubit is $|-\theta\rangle$. Four of the outcomes involve learning the result of the first or second qubit. The other two options involve learning the XOR of the bits. For example if Bob obtains outcome F then he learns the two qubits are different and so the XOR of them is 1. Similarly for the rest of the results,

$$\begin{aligned} A =& \{|\theta,\theta\rangle, |\theta,-\theta\rangle\}, && \text{1st qubit is } |-\theta\rangle \\ B =& \{|\theta,\theta\rangle, |-\theta,\theta\rangle\}, && \text{2nd qubit is } |-\theta\rangle \\ C =& \{|\theta,-\theta\rangle, |-\theta,-\theta\rangle\}, && \text{2nd qubit is } |\theta\rangle \\ D =& \{|-\theta,\theta\rangle, |-\theta,-\theta\rangle\}, && \text{1st qubit is } |\theta\rangle \\ E =& \{|\theta,-\theta\rangle, |-\theta,\theta\rangle\}, && \text{qubits are the same } XOR = 0 \\ F =& \{|\theta,\theta\rangle, |-\theta,-\theta\rangle\}, && \text{qubits are different } XOR = 1. \end{aligned} \tag{3.101}$$

For the outcomes where we learn either the first or second qubit we equate $|\theta\rangle \rightarrow 0$ and $|-\theta\rangle \rightarrow 1$ and take the known bit to be in the secret shared bit string. For the other two scenarios we just take the XOR as the bit value. So there are three outcomes for it being 0 (C, D, E), and three outcomes for it being 1 (A, B, F) and with a random input this will produce a random string. Also looking at the probabilities of each outcome from equation (3.85) we can see that there is an equal chance of the bit being 0 or 1.

After the measurement Bob announces publicly whether he obtained the first bit, second bit or the XOR and then Alice looks at what she sent and from that a secret string can be formed. In this scenario every single sent state will be used in the final string and Alice is not required to publicly declare her measurement basis. The downside is that two qubits are required for each state. This deterministic factor that doesn't occur in BB84 is

also available in a similar QKD protocol based upon a photon version of the mean king problem [43].

Security against simple attacks comes from a similar method as in BB84, due to the fact any interference from Eve will alter the states between Alice and Bob. If Bob and Alice publicly check a portion of the shared key for any errors, then if errors exist it is known Eve has attempted to measure the states and the key will be discarded and the process started again. There has been no in depth security proof yet and neither have we analysed it further yet to see if it has any improvements on current QKD protocols. This protocol is similar to B92 in that the only sifting required is an announcement from Bob. Susceptibility of the practical setup is potentially an issue here as first of all the requirement for product states and an entangled measurement basis could be technically more demanding and inefficiencies due to imperfect sources, detectors and other optical elements would increase the susceptibility of the protocol to attacks. There are many options to counter these problems for example using weak coherent pulses or decoy states[57] to reduce the effect of a photon number splitting attack. This is definitely something to look at in the future to determine the feasibility of this two out of four measurement as a QKD protocol.

3.5 Average Number Of States Eliminated

A different figure of merit to investigate is attempting to eliminate as many states as possible on average. Again we will be looking at states in the form of $|\pm\theta, \pm\theta....\rangle$, where we use N to describe the number of qubits. We shall investigate whether it will be best to perform measurements on the single qubits or if a global measurement will be optimal.

3.5.1 Individual Measurements

If we choose to measure each qubit individually then we can use the binomial distribution to give the average number of states eliminated. The binomial distribution gives the probability of obtaining n successes from N trials with a success probability p_s and the failure probability $p_f = 1 - p_s$. This probability is given by

$$P(n|N) = \binom{N}{n} p_s^n \, p_f^{N-n}, \tag{3.102}$$

where $\binom{N}{n}$ is the binomial coefficient.

As we are looking for the average number of states eliminated we need to weight each measurement outcome with the number of states they will eliminate. In our case a successful measurement on an individual qubit will eliminate half of the remaining states. In

the table below we show how many states are eliminated (N_e) for the number of successful measurements (n) for N qubits.

n	N_e
0	0
1	$2^N/2$
2	$2^N/2 + 2^N/4$
$\vdots$	$\vdots$
n	$2^N - \sum_{n=0}^{N} 2^N/2^n$.

(3.103)

Alternatively you can look at the number of failed measurements (m), then the number of states remaining (N_r) is simply 2^m. Using this simpler approach we can calculate the average number of states remaining as

$$\langle N_r \rangle = \sum_{m=0}^{N} \binom{N}{m} p_f^m (1-p_f)^{(N-m)} 2^m. \tag{3.104}$$

To calculate the number of states eliminated we simply need to subtract this from the total number of states,

$$\begin{aligned} \langle N_e \rangle &= 2^N - \langle N_r \rangle = 2^N - \sum_{m=0}^{N} \binom{N}{m} p_f^m (1-p_f)^{(N-m)} 2^m \\ &= 2^N - [2p_f + (1-p_f)]^N = 2^N - (1+p_f)^N. \end{aligned} \tag{3.105}$$

If we are looking at unambiguous measurements then the success probability for unambiguous discrimination between the two states $|\theta\rangle$ and $|-\theta\rangle$ defined in equation (1.44) is given by, $p_f = |\langle -\theta | +\theta \rangle| = \cos(2\theta)$. This is the success probability for each individual measurement on an individual qubit.

This can be checked looking at the two-qubit state with a separation angle $\pi/4$, which equates to $\theta = \pi/8$, which is the same overlap as in the PBR measurement. We can eliminate zero, two or three states with individual measurements. Each individual measurement has two possible outcomes, success (S) or fail (F).

States eliminated	Outcomes
3	SS
2	SF or FS
0	FF

.

$$\begin{aligned} \langle N_e \rangle &= 3p_s^2 + 2(2p_s(1-p_s)) + 0(1-p_s)^2 \\ &= 3(1-p_f)^2 + 2(2p_f(1-p_f)) + 0(p_f^2) \\ &= 3(1+p_f^2-2p_f) + 4p_f - 4p_f^2 \\ &= 3 - 2p_f - p_f^2 \\ &= 4 - (1+p_f)^2. \end{aligned} \tag{3.106}$$

This is equivalent to the result from equation (3.105) with $N = 2$. Now if we substitute in $p_f = \cos(2\pi/8) = \sqrt{2}/2$ to create the same overlaps as in the PBR measurement we get

$$\langle N_e \rangle = 4 - \left(1 + \frac{\sqrt{2}}{2}\right)^2 = 1.086. \tag{3.107}$$

Measurements on individual qubits eliminate on average 1.086 states, which is more than the PBR measurement, which eliminates one state. On the other hand, the PBR measurement is guaranteed to always eliminate exactly one two-qubit state, while individual measurements sometimes fail to eliminate any state at all.

3.5.2 Upper Bounds For Elimination Measurements

By using the proven optimal results for single state unambiguous discrimination derived by Ivanovic, Dieks and Peres (IDP) [11][12][13] we can put an upper bound on the success probabilities for the elimination measurements so that they do not perform unambiguous state discrimination in a manner that outperforms the IDP bound. For all the proofs below we use the states $|0\rangle$ and $|+\rangle$, yet the proof holds with any two different states. The values of the success and failure probabilities will vary but these are not specified in the proofs. This is the same concept as used to get the bound in the optimal result for unambiguously eliminating two out of four states but here we have a more general result.

Two Qubits

For any initial state we have six possible results. If we say the initial state was $|00\rangle$ then the possible results can't include eliminating $|00\rangle$. In the case the initial state was $|00\rangle$ then the following six outcomes are possible, where the number on the left hand side denotes the number of states eliminated and on the right we have the combinations of states eliminated. These are denoted by a letter that we will refer to in the rest of this section.

$$\begin{aligned}
&3: \quad G = \{|0+\rangle, |+0\rangle, |++\rangle\}, \\
&2: \quad D = \{|0+\rangle, |+0\rangle\},\ E = \{|0+\rangle, |++\rangle\},\ F = \{|+0\rangle, |++\rangle\}, \\
&1: \quad A = \{|0+\rangle\} \qquad B = \{|+0\rangle\} \qquad C = \{|++\rangle\}.
\end{aligned} \tag{3.108}$$

To bound the success probabilities we will use the assumptions we know one of the qubits or some information about the qubits and then with this we check that the elimination measurements can't outperform the IDP bound for unambiguous state discrimination. I will use the notation of $|?\rangle$ to describe a qubit we don't know. For example if we know the first qubit is $|0\rangle$ and the second qubit is unknown this will be written as $|0?\rangle$. Then if the initial state was $|00\rangle$ we know we can't eliminate any set of states containing $|00\rangle$

and also any measurement that involves eliminating $|0+\rangle$ will tell us the second qubit is $|0\rangle$. Therefore the probability to eliminate a set containing $|0+\rangle$ (A,D,E or G) can't exceed the probability to discriminate between $|0\rangle$ and $|+\rangle$. Now we will go through some of the limitations that occur from not allowing elimination measurements to perform discrimination better than the IDP bound assuming the initial state was $|00\rangle$.

- If we know the first qubit $|0?\rangle$, then the outcomes that tell us what the second qubit is (eliminate $|0+\rangle$) are A, D, E and G. Therefore the probability of eliminating set A, D, E or G can't exceed the probability to discriminate between $|0\rangle$ and $|+\rangle$. This gives us the bound $p_A + p_D + p_E + p_G \leq 1 - p_f$, where p_f is the failure probability, which in our approach is given by the IDP bound.

- If we know the second qubit $|?0\rangle$, then the outcomes that tell us what the second qubit is (eliminate $|+0\rangle$) are B, D, F and G.

- If we know that the state is either $|00\rangle$ or $|++\rangle$, then the measurement outcomes that tell us it is $|00\rangle$ (eliminate $|++\rangle$) are C, E, F and G

Each of these requirements give us the following bounds respectively,

$$\begin{aligned} p_A + p_D + p_E + p_G &\leq 1 - p_f, \\ p_B + p_D + p_F + p_G &\leq 1 - p_f, \\ p_C + p_E + p_F + p_G &\leq 1 - p_f^2. \end{aligned} \tag{3.109}$$

The top two bounds have a single power of p_f as we are just discriminating a single qubit, whereas in the third scenario we are discriminating both the first and second qubit so we have the $1 - p_f^2$ term. If we sum up the three inequalities above then we get

$$S_2 = 3p_G + 2(p_D + p_E + p_F) + p_A + p_B + p_C \leq 3 - 2p_f - p_f^2. \tag{3.110}$$

Coincidently the left hand side is the average number of states eliminated. This is because the probabilities are weighted by how many states they eliminate. G eliminates three states, D, E, F eliminate two and A, B, C eliminate one state. Now this has given us a bound on the average states eliminated. We have formed this bound assuming the initial state was $|00\rangle$, the result would be the same if we took any initial state. This is because each of the states has the same symmetry about the other three states. The right hand side of the bound in equation (3.110) also happens to be the number of states eliminated by two unambiguous individual measurements as can be shown below,

$$3(1 - p_f)^2 + 2p_f(1 - p_f) = 3 - 2p_f - p_f^2. \tag{3.111}$$

Therefore we have also shown that in two-qubit case a method to eliminate the highest average number of states is by individual measurements on each qubit as that would saturate the bound given in (3.110).

Three Qubits

Finding out that individual unambiguous discrimination measurements was an optimal way to eliminate the highest average number of states unambiguously was an interesting outcome in the two-qubit case. I will now go on to check this for three qubits to see if we obtain the same result.
If individual measurement were the optimal method in the three-qubit (8 state) scenario we would expect the bound to be,

$$7(1-p_f)^3 + 6 \times 3(1-p_f)^2 p_f + 4 \times 3(1-p_f)p_f^2 \leq 7 - 3p_f - 3p_f^2 - p_f^3. \quad (3.112)$$

The derivation will follow as of that for two qubits but after defining the states that will eliminate one of the eight options with a letter I will just use that letter instead of explicitly writing out all the states each time. Also for the measurements that eliminate more than half of the states I will show it by describing what is not eliminated thus reducing the number of states to be expressed. For example say $A1$ denotes eliminating seven of the eight states and $G1$ denotes eliminating $|00+\rangle$ then $A1 - G1$ will eliminate the other six states in $A1$ that are not $|00+\rangle$.

For this approach we assume the sent state was $|000\rangle$. In the following section the number on the left describes how many states are being eliminated and then to the right we have

the different possible ways to eliminate that number of states.

$7:\quad A1 = \{|00+\rangle, |0+0\rangle, |0++\rangle\, |+00\rangle, |+0+\rangle, |++0\rangle, |+++\rangle\}.$

$1:\quad G1 = |00+\rangle,\ G2 = |0+0\rangle,\ G3 = |0++\rangle,\ G4 = |+00\rangle,\ G5 = |+0+\rangle,$
$G6 = |++0\rangle,\ G7 = |+++\rangle.$

$2:\quad F1 = G1G2,\ F2 = G1G3,\ F3 = G1G4,\ F4 = G1G5,\ F5 = G1G6,$
$F6 = G1G7,\ F7 = G2G3,\ F8 = G2G4,\ F9 = G2G5,\ F10 = G2G6,$
$F11 = G2G7,\ F12 = G3G4,\ F13 = G3G5,\ F14 = G3G6,\ F15 = G3G7,$
$F16 = G4G5,\ F17 = G4G6,\ F18 = G4G7,\ F19 = G5G6,\ F20 = G5G7,$
$F21 = G6G7.$

$3:\quad E1 = G1G2G3,\ E2 = G1G2G4,\ E3 = G1G2G5,\ E4 = G1G2G6,$
$E5 = G1G2G7,\ E6 = G1G3G4,\ E7 = G1G3G5,\ E8 = G1G3G6,$
$E9 = G1G3G7,\ E10 = G1G4G5,\ E11 = G1G4G6,\ E12 = G1G4G7,$
$E13 = G1G5G6,\ E14 = G1G5G7,\ E15 = G1G6G7,\ E16 = G2G3G4,$
$E17 = G2G3G5,\ E18 = G2G3G6,\ E19 = G2G3G7,\ E20 = G2G4G5,$
$E21 = G2G4G6,\ E22 = G2G4G7,\ E23 = G2G5G6,\ E24 = G2G5G7,$
$E25 = G2G6G7,\ E26 = G3G4G5,\ E27 = G3G4G6,\ E28 = G3G4G7,$
$E29 = G3G5G6,\ E30 = G3G5G7,\ E31 = G3G6G7,\ E32 = G4G5G6,$
$E33 = G4G5G7,\ E34 = G4G6G7,\ E35 = G5G6G7.$

$4:\quad D1 = A1 - G1G2G3,\ D2 = A1 - G1G2G4,\ D3 = A1 - G1G2G5,$
$D4 = A1 - G1G2G6,\ D5 = A1 - G1G2G7,\ D6 = A1 - G1G3G4,$
$D7 = A1 - G1G3G5,\ D8 = A1 - G1G3G6,\ D9 = A1 - G1G3G7,$
$D10 = A1 - G1G4G5,\ D11 = A1 - G1G4G6,\ D12 = A1 - G1G4G7,$
$D13 = A1 - G1G5G6,\ D14 = A1 - G1G5G7,\ D15 = A1 - G1G6G7,$
$D16 = A1 - G2G3G4, D17 = A1 - G2G3G5,\ D18 = A1 - G2G3G6,$
$D19 = A1 - G2G3G7,\ D20 = A1 - G2G4G5,\ D21 = A1 - G2G4G6,$
$D22 = A1 - G2G4G7,\ D23 = A1 - G2G5G6,\ D24 = A1 - G2G5G7,$
$D25 = A1 - G2G6G7,\ D26 = A1 - G3G4G5,\ D27 = A1 - G3G4G6,$
$D28 = A1 - G3G4G7,\ D29 = A1 - G3G5G6,\ D30 = A1 - G3G5G7,$
$D31 = A1 - G3G6G7,\ D32 = A1 - G4G5G6,\ D33 = A1 - G4G5G7,$
$D34 = A1 - G4G6G7,\ D35 = A1 - G5G6G7.$

$$
\begin{aligned}
5: \quad & C1 = A1 - G1G2,\ C2 = A1 - G1G3,\ C3 = A1 - G1G4,\ C4 = A1 - G1G5, \\
& C5 = A1 - G1G6,\ C6 = A1 - G1G7,\ C7 = A1 - G2G3,\ C8 = A1 - G2G4, \\
& C9 = A1 - G2G5,\ C10 = A1 - G2G6,\ C11 = A1 - G2G7,\ C12 = A1 - G3G4, \\
& C13 = A1 - G3G5,\ C14 = A1 - G3G6,\ C15 = A1 - G3G7,\ C16 = A1 - G4G5, \\
& C17 = A1 - G4G6,\ C18 = A1 - G4G7,\ C19 = A1 - G5G6,\ C20 = A1 - G5G7, \\
& C21 = A1 - G6G7. \\
6: \quad & B1 = A1 - G1,\ B2 = A1 - G2,\ B3 = A1 - G3,\ B4 = A1 - G4, \\
& B5 = A1 - G5,\ B6 = A1 - G6,\ B7 = A1 - G7. \qquad (3.113)
\end{aligned}
$$

In the following section we go through the scenarios in which these elimination outcomes could break the individual qubit unambiguous discrimination bound and then limit these using the IDP bound, thus creating a bound on the elimination of states. Each scenario has a label i associated to it and then the probability for those scenarios is described by p_i. The question mark refers to a case when we don't know what that qubit is. Like in the two-qubit case we assume an initial state and in this case it will be $|000\rangle$ so none of the measurements can involve eliminating $|000\rangle$.

1. If we know the first two qubits so the state is $|00?\rangle$, then the only two options left are $|000\rangle$ and $|00+\rangle$. As the initial state was $|000\rangle$ to learn the third qubit need to eliminate $|00+\rangle$ (G1). In the list below are all the possible elimination combinations that include G1 from the list in (3.113). The bound on this is $p_1 \leq 1 - p_f$ as we are just discriminating a single qubit (the third qubit).

A	1
B	2-7
C	7-21
D	16-35
E	1-15
F	1-6
G	1

2. If we know the first and last qubit so the state is $|0?0\rangle$, to learn the middle qubit we need to eliminate G2 ($|0+0\rangle$). The bound on this is $p_2 \leq 1 - p_f$. Below are all the measurement outcomes containing G2,

A	1
B	1,3-7
C	2-6,12-21
D	6-15, 26-35
E	1-5, 16-25
F	1, 7-11
G	2

3. If we know the second and third qubit so the state is $|?00\rangle$, to learn first qubit we need to eliminate G4($|+00\rangle$). The bound on this is $p_3 \leq 1 - p_f$. Below are all the measurement outcomes containing G4,

A	1
B	1-3,5-7
C	1,2,4-7,9-11,13-15,19-21
D	1,3-5,7-9,13-15,17-19,23-25,29-31,35
E	2,6,10-12,16,20-22,26-28,32-34
F	3,8,12,16-18
G	4

4. If we know only the first qubit $|0??\rangle$ then to eliminate G3 ($|0++\rangle$) we would require at least one successful individual measurement on the second or third qubit. For one successful measurement out of two chances the probability is just one minus the probability of failing both times ($1 - p_f^2$), so the bound is $p_4 \leq 1 - p_f^2$. Below are all the measurement outcomes containing G3,

A	1
B	1,2,4-7
C	1,3-6,8-11,16-21
D	2-5,10-15,20-25,32-35
E	1, 6-9, 16-19, 26-31
F	2,7,12-15
G	3

5. If we know $|?0?\rangle$ then to eliminate G5 $|+0+\rangle$) we again require at least one successful measurement out of two on the first and third qubit this time, giving us the bound $p_5 \leq 1 - p_f^2$. Below are all the measurement outcomes containing G5,

A	1
B	1-4,6,7
C	1-3,5-8,10-12,14,15,17-18,21
D	1,2,4-6,8,9,11,12,15,16,18,19,21,22,25,27,28,31,34
E	3,7,10,13,14,17,20,23,24,26,29,30,32,33,35
F	4,9,13,16,19,20
G	5

6. If we know $|0??\rangle$ then to eliminate G6 $|{+}{+}0\rangle$) we require at least one successful measurement out of two on the first or second qubit, giving us the bound $p_6 \leq 1 - p_f^2$. Below are all the measurement outcomes containing G6,

A	1
B	1-5,7
C	1-4,6-9,11-13,15,16,18
D	1-3,5-7,9,10,12,14,16,17,19,20,22,24,26,28,30,33
E	4,8,11,13,15,18,21,23,25,27,29,31,32,34,35
F	5,10,14,17,19-21
G	6

7. Tell us it's $|000\rangle$ if we know it is either $|000\rangle$ or $|{+}{+}{+}\rangle$ then to eliminate G7 we require at least one successful measurement out of three. $p_7 \leq 1 - p_f^3$. Below are all the measurement outcomes containing G7,

A	1
B	1-6
C	1-5,7-10,12-14,16-17,19
D	1-4,6-8,10-11,13,16-18,20-21,23,26-27,29,32
E	5,9,12,14,15,19,22,24,25,30,31,33-35
F	6,11,15,18,20-21
G	7

The right hand side of our bound will involve summing up all the inequalities given by the seven p_i's above, that arise from the limitations due to the IDP bound.

$$\begin{aligned}\sum_{i=1}^{7} p_i &= 3(1-p_f) + 3(1-p_f^2) + (1-p_f^3) \\ &= 7 - 3p_f - 3p_f^2 - p_f^3. \end{aligned} \tag{3.114}$$

This is the same as in equation (3.112) so now we just need to count up the eliminated states from the seven scenarios above to see if we obtain the left hand side of (3.112). For

average states we are expecting all the G terms to appear once as they eliminate one state, F terms twice as they eliminate two states and so on so forth down to the B states six times and $A1$ seven. So adding up all the times we obtain these outcomes from the seven scenarios in the lists above we get,

- For the G states we have $G1 + G2 + G3 + G4 + G5 + G6 + G7$ as each appears once in each of the seven scenarios.
- For the B states we have 6(B1+B2+B3+B4+B5+B6+B7) as each of the seven scenarios has all but one of the B states giving us six of each B states.
- For the C states we end up with five of all C states and so therefore there must be two of all the F states due to the symmetry in the definition of the outcomes.
- For D and E we have four and three sets of each outcome respectively.

Therefore we have again shown that the optimal measurement for eliminating the highest average number of states is that of individual unambiguous measurements on each qubit, this time in the three-qubit scenario. As we have a result for two and three qubits it is likely this extends to any number of qubits so we shall introduce a general result in the next section.

N Qubits

As we have a proof for two and three qubits we can use a similar approach for N qubits. I will try and explain it with examples from the above two cases to help verify the points. First let us say the state is going to be $|0\rangle^{\otimes N}$. Now we take another N-qubit state $|x\rangle$ that in each qubit position has either a $|0\rangle$ or a $|+\rangle$. If we wish to determine whether $|0\rangle^{\otimes N}$ or $|x\rangle$ was sent the probability of this can not exceed $1 - p_f^M$, where M denotes the number of qubit positions in $|0\rangle^{\otimes N}$ and $|x\rangle$ that are different. In information theory this is referred to as the Hamming distance This is because any successful measurement on an individual qubit in a position where $|0\rangle^{\otimes N}$ and $|x\rangle$ are different will allow us to discriminate between the two. With this we can give the probability of ruling out $|x\rangle$ as,

$$p(\neg\, x) \leq 1 - p_f^M. \tag{3.115}$$

There are $2^N - 1$ ways to choose $|x\rangle$ and within this there are $\binom{N}{M}$ different ways for a certain M. This means that if $|0\rangle^{\otimes N}$ and $|x\rangle$ differ in three places ($M = 3$) there are $\binom{N}{3}$ ways the states $|0\rangle^{\otimes N}$ and $|x\rangle$ can differ. Analogous to the two-qubit and three-qubit bounds given in (3.110) and (3.112) respectively the right hand side of the bound for N qubits is

$$\sum_{M=0}^{N} \left[\binom{N}{M} (1 - p_f^M) \right]. \tag{3.116}$$

This can be simplified as,

$$\sum_{M=0}^{N} \binom{N}{M} = 2^N \qquad \text{and} \qquad \sum_{M=0}^{N} \binom{N}{M} p_f^M = (1+p_f)^N, \tag{3.117}$$

so therefore,

$$\sum_{M=0}^{N} \left[\binom{N}{M} (1-p_f^M) \right] = 2^N - (1+p_f)^N. \tag{3.118}$$

Now we have the same right hand side as that from which we would achieve using individual unambiguous measurements on each qubit as shown in equation (3.105). Now we need to collect all the probabilities of eliminating a certain state $|x\rangle$ to form the other side of our bound.

Firstly $p(\neg\, x)$ from equation (3.115) needs a contribution of how many states it eliminates, this being anywhere from 1 to $2^N - 1$. If we introduce the probability $p(\neg\, K)$, where K is the number of states eliminated, there are $\binom{2^N-1}{K}$ ways of eliminating K states and of these $\binom{2^N-2}{K-1}$ obtain $|x\rangle$. As $|x\rangle$ must be included in the eliminated states. Therefore the coefficient of $p(\neg\, K)$ is:

$$(2^N-1)\frac{\binom{2^N-2}{K-1}}{\binom{2^N-1}{K}} = \frac{(2^N-1)!K!(2^N-1-K)!}{(K-1)!(2^N-1-K)!(2^N-1)!} = K. \tag{3.119}$$

So the sum for eliminating all $2^N - 1$ numbers of states can be described as:

$$\sum_{K=1}^{2^N-1} Kp(\neg K) = 2^N - (1+p_f)^N. \tag{3.120}$$

The left hand side is the average number of states eliminated and so we have a general proof for N qubits that individual unambiguous measurements are the best approach for eliminating the highest average number of states.

3.5.3 SDP Results Compared Against Local Measurements

From semidefinite programming we can approximately find the minimum angles we require to eliminate M states, where the angle 2θ is the angle between the two possible states for a qubit in each position. One can then compare the number of states eliminated to the average number of states that can be eliminated by individual qubit measurements (local) at the limiting angles from semidefinite programming. For local unambiguous measurements we have the average number eliminated as

$$\langle N_{elocal} \rangle = 2^N - (1+p_f)^2 = 2^N - (1+\cos(2\theta))^2. \tag{3.121}$$

Two Qubits

In the following table I have the minimum angle found by SDP to eliminate the respective number of states conclusively (probability of success is one). Then the third column gives the average number of states eliminated using unambiguous discrimination measurements upon each qubit ($\langle N_{elocal} \rangle$).

Number of States Eliminated	Minimum Angle	$\langle N_{elocal} \rangle$
1	$22.5°$	1.086
2	$32.765°$	2
3	$45°$	3

For eliminating all but one state the best approach is individual measurements and so we know that the average states will be the same. We can also see that at the angle required to eliminate one state with certainty it is possible to eliminate more than one state on average. Interestingly for at $\theta = 32.765°$ the local measurement can't eliminate more on average than the approach that eliminates two out of four with certainty every time, even though the measurement we found to eliminate two of the four states was not individual unambiguous measurements on each qubit.

Three Qubits

The following table shows the minimum angle from SDP to eliminate the respective number of states conclusively compared to the number from local measurements upon each qubit.

Number of States Eliminated	Minimum Angle	$\langle N_{elocal} \rangle$
1	$14.570°$	1.42
2	$18.284°$	2.14
3	$22.725°$	3.07
4	$27.014°$	4.00
5	$32.012°$	5.02
6	$37.467°$	6.00
7	$45°$	7

For three qubits there are some precision issues with the SDP and finding the exact angle of conclusive elimination is not particularly easy. Generally it seems though as you try and eliminate more states the conclusive measurement's ability becomes similar to that of the local approach at eliminating the highest average number. This is most likely due to the fact the measurement itself is becoming more local as we know to eliminate all but one it is a completely local measurement.

3.6 Conclusion

In this chapter we introduced a collection of elimination measurements for certain scenarios. We predominantly looked at the two-qubit case akin to the PBR measurement finding the optimal measurement for eliminating two out of out the four two-qubit states unambiguously. We found what we believe to be an optimal measurement for eliminating one and two out of the four states unambiguously. From this we introduced a potential use of this measurement for cryptography. To extend this work it would be interesting to look at measurements for longer sequences of qubits and see if there was some general term for the probability of successful measurement for the elimination of m out of the 2^N possible states. A security analysis of the the QKD protocol would be interesting to see how it performs compared to the current protocols and whether it was practically feasible.

In the next chapter we look into the method of decomposing a unitary into beamsplitter like operations and therefore forming an experimental setup for any measurement. We apply this to our two out of four elimination measurement as well as looking at the process in a more general manner.

Chapter 4

Decomposition of Unitary Matrices into Optical Elements

A quantum measurement can be represented by a unitary matrix and it was shown by Reck et al. [58] that any discrete finite-dimensional unitary operator can be constructed in the laboratory using optical devices. This is done by decomposing a square unitary matrix U, into a sequence of two-dimensional transformations that can be represented by an optical element, for example a beam splitter or phase shifter. It is not a requirement to use optical devices, it is only a requirement that the two-dimensional transforms are mathematically equivalent to a beamsplitter like operation. They could for example be a laser pulse coupling two atomic levels.

4.1 Neumark Extension

The decomposition works for any square unitary matrix and therefore you will be able to represent a projective von Neumann-measurement [5] with the rows as the states you are projecting onto and the columns as the basis states. For a generalised measurement the number of measurement outcomes is not limited by the dimensions of the problem. The Neumark (or Naimark) extension [6] is a way to realise any generalised measurement as a projective von Neumann-measurement in an extended higher-dimensional space. This can be done by coupling the system to an auxiliary basis,. For example if we have an initial system in the state $|\psi\rangle_S$, and then this can be coupled to an auxiliary state $|\phi\rangle_{aux}$ using the unitary transform $\hat{U}|\psi\rangle_S \otimes |\phi\rangle_{aux}$. If we have a generalised measurement with M rank one operators Π_i then the M pure states can be written as

$$|\psi_i\rangle = \sum_j a_{ij}|j\rangle, \tag{4.1}$$

where $j = 1, 2, ..., N$. Then we can organise the a_{ij} coefficients as an M x N matrix,

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{pmatrix}, \tag{4.2}$$

where the M rows correspond to the states $|\psi_i\rangle$, and the N columns are the basis states. As the columns are the basis states they are orthonormal, therefore to form a square unitary matrix we can add the $M - N$ columns by choosing any set of vectors that hold the orthonormal requirement. Once the conditions are satisfied there will still be an infinite amount of choice up until the last column which will be unique up to a phase factor. These $M - N$ columns represent the auxiliary basis states.

4.2 Decomposition of a unitary

Once you have a square unitary matrix then we can decompose the matrix using the following method. First of all we define our general 2 x 2 unitary transform as

$$U(2) = e^{i\phi} \begin{pmatrix} a & -b \\ b^* & a^* \end{pmatrix}, \tag{4.3}$$

where $|a|^2 + |b|^2 = 1$ and $e^{i\phi}$ is a global phase shift. Let us also define a matrix T_{ij} that is an identity matrix of the dimension of the original measurement unitary, with the $U(2)$ matrix from equation (4.3) in the ij positions. For example if the unitary matrix you are decomposing is 6 x 6 then T_{42} would be

$$T_{42} = e^{i\phi} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & a & 0 & -b & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & b^* & 0 & a^* & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.4}$$

To decompose the unitary U_N we are aiming to diagonalise the matrix by making every element under the diagonal zero. This is done by multiplying the unitary with a sequence of T_{ij} matrices starting from $T_{N,N-1}$, then decreasing the column label j of T_{ij} through $N-1, N-2, ...1$. To calculate the composition of each T_{ij} matrix we multiply U_6 by T_{ij} then we solve the for the targeted element so that $U_{ij} = 0$, where U_{ij} is the i, j element of U. Setting the element U_{ij} as zero gives us an equation in the form $f(a, b) = 0$, then we solve this along with the condition $|a|^2 + |b|^2 = 1$ to get our values of a and b to substitute

back in to the T_{ij} matrix. After the reduction of the bottom row you get,

$$U_N.T_{N.N-1}.T_{N,N-2}...T_{N,1} = \left(\begin{array}{ccccc|c} & & U_{N-1} & & & 0 \\ & & & & & 0 \\ & & & & & \vdots \\ & & & & & 0 \\ \hline 0 & 0 & \dots & & 0 & e^{i\phi_N} \end{array}\right). \tag{4.5}$$

Once the bottom row (excluding the diagonal) has been turned to zeroes then we move up a row so reduce the i label of T_ij by one. We repeat the process except this time we only act on the columns $N-2, N-3, ..1$ as we don't wish to act on the diagonal elements of U_N or anywhere to the right of the diagonal elements. This is because as it is a unitary matrix once we make $U_{ij} = 0$, then this means $U_{ji} = 0$. This is then repeated all the way until each non-diagonal element under the diagonal of the matrix has been multiplied by it's respective T_{ij} and become zero. The final decomposed matrix is then,

$$U_N.T_{N.N-1}.T_{N,N-2}...T_{2,1} = D = \begin{pmatrix} e^{i\phi_1} & 0 & \dots & 0 \\ 0 & e^{i\phi_2} & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\phi_N} \end{pmatrix}. \tag{4.6}$$

We can therefore also reconstruct the matrix U_N out of the T matrices and U_N can be written as

$$U_N = DT^\dagger_{2,1}...T^\dagger_{N,N-2}.T^\dagger_{N,N-1}. \tag{4.7}$$

This shows that we can realise any unitary as a sequence of beam splitter type operations. The matrix D at the end is just the identity with phase shifts upon each mode. If for example we used spatial modes with photons in each mode then D would represent the final phase for each photon in each mode. If we just detect the presence of a photon to show which mode was the final outcome then the phase is irrelevant so in this scenario the phases of D don't matter and the result would be the same if D was the identity.

4.3 Unambiguous State Discrimination

As an example we will study how to form a unitary and then decompose said unitary for unambiguous state discrimination. Our aim will be to unambiguously discriminate between

$$|\theta\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle \qquad \text{and} \qquad |-\theta\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle. \tag{4.8}$$

The three measurement outcomes will be eliminating $|\theta\rangle, |-\theta\rangle$ and an outcome corresponding to the measurement failing to eliminate either state. The measurement operators can be written as,

$$\begin{aligned}
\Pi_{\overline{\theta}} &= a_0|\overline{\theta}\rangle\langle\overline{\theta}| = a_0(\sin^2\theta|0\rangle\langle 0| + \cos^2\theta|1\rangle\langle 1| - \cos\theta\sin\theta(|0\rangle\langle 1| + |1\rangle\langle 0|)),\\
\Pi_{\overline{-\theta}} &= a_1|\overline{-\theta}\rangle\langle\overline{-\theta}| = a_1(\sin^2\theta|0\rangle\langle 0| + \cos^2\theta|1\rangle\langle 1| + \cos\theta\sin\theta(|0\rangle\langle 1| + |1\rangle\langle 0|)),\\
\Pi_f &= I - (\Pi_{\overline{\theta}} + \Pi_{\overline{-\theta}}) = (1 - (a_0 + a_1)\sin^2\theta)|0\rangle\langle 0| + (1 - (a_0 + a_1)\cos^2\theta)|1\rangle\langle 1|\\
&\qquad + (a_0 - a_1)\cos\theta\sin\theta(|01\rangle\langle 01| + |10\rangle\langle 10|), \quad (4.9)
\end{aligned}$$

where $|\overline{\theta}\rangle$ and $|\overline{-\theta}\rangle$ are the orthogonal states given by

$$|\overline{\theta}\rangle = (\sin\theta|0\rangle - \cos\theta|1\rangle) \qquad \text{and} \qquad |\overline{-\theta}\rangle = (\sin\theta|0\rangle + \cos\theta|1\rangle), \tag{4.10}$$

I is the identity matrix and a_0, a_1 are the weighting of each projector. For equal a priori probabilities between state $|\theta\rangle$ and $|-\theta\rangle$ and the same weighting upon each projector $(a_0 = a_1)$ we have

$$\Pi_f = I - (\Pi_{\overline{\theta}} + \Pi_{\overline{-\theta}}) = (1 - 2a_0\sin^2\theta)|0\rangle\langle 0| + (1 - 2a_0\cos^2\theta)|1\rangle\langle 1|. \tag{4.11}$$

Π_f can't immediately be written as a single rank one measurement operator in this case, but we can make it up of two rank one operators Π_{f1} and Π_{f2}, which can be written as

$$\begin{aligned}
\Pi_{f1} &= |\psi_1\rangle\langle\psi_1|, \qquad \text{where} \qquad |\psi_1\rangle = \sqrt{1 - 2a_0\sin^2\theta}|0\rangle,\\
\Pi_{f2} &= |\psi_2\rangle\langle\psi_2|, \qquad \text{where} \qquad |\psi_2\rangle = \sqrt{1 - 2a_0\cos^2\theta}|1\rangle. \quad (4.12)
\end{aligned}$$

An outcome associated with Π_{f1} or Π_{f2} will be denoted as a failure outcome. For equal a priori probabilities the optimum measurement was shown to have

$$a_0 = a_1 = \frac{1}{2\cos^2\theta}, \tag{4.13}$$

and this makes $\Pi_{f2} = 0$ so our failure operator can now be written as a rank one operator given by $\Pi_f = 1 - \tan^2\theta|0\rangle\langle 0|$. If the prior probabilities are not the same then the optimal measurement [14] has

$$a_0 = \frac{1 - \sqrt{\frac{p_{-\theta}}{p_\theta}}\cos(2\theta)}{\sin^2(2\theta)} \qquad \text{and} \qquad a_1 = \frac{1 - \sqrt{\frac{p_\theta}{p_{-\theta}}}\cos(2\theta)}{\sin^2(2\theta)}, \tag{4.14}$$

where p_θ and $p_{-\theta}$ are the probability of states $|\theta\rangle$ and $|-\theta\rangle$ occurring respectively. For equal probabilities though if we substitute in the values of a_0 and a_1 from (4.13), we can derive the pure states of our projector from our M x N matrix in (4.2) as,

$$\begin{pmatrix} \frac{\tan\theta}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{\tan\theta}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \sqrt{1 - \tan^2\theta} & 0 \end{pmatrix}. \tag{4.15}$$

Introducing a third basis state that we shall call the auxiliary basis state $|aux\rangle$ to make it a square matrix, we have the 3 X 3 matrix,

$$\begin{pmatrix} \frac{\tan\theta}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \alpha \\ \frac{\tan\theta}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \beta \\ \sqrt{1-\tan^2\theta} & 0 & \gamma \end{pmatrix}. \tag{4.16}$$

The requirement for this basis state is that it is orthonormal to the current basis states. This gives us the set of requirements,

$$\frac{\alpha\tan\theta}{\sqrt{2}} + \frac{\beta\tan\theta}{\sqrt{2}} + \gamma\sqrt{1-\tan^2}\theta = 0, \tag{4.17}$$

$$\frac{\alpha}{\sqrt{2}} - \frac{\beta}{\sqrt{2}} = 0, \tag{4.18}$$

$$\alpha^2 + \beta^2 + \gamma^2 = 1. \tag{4.19}$$

From (4.18) we can see $\alpha = \beta$. Re-arranging (4.17) we get,

$$\sqrt{2}\alpha\tan\theta + \gamma\sqrt{1-\tan^2\theta} \;\rightarrow\; \gamma = -\frac{\sqrt{2}\alpha\tan\theta}{\sqrt{1-\tan^2\theta}}. \tag{4.20}$$

Substituting these expression for β and γ into equation (4.19) we just have to solve

$$2\alpha^2 + \frac{2\alpha^2\tan^2\theta}{1-\tan^2\theta} = 1, \tag{4.21}$$

for α. Taking the positive solutions to the square roots as our final answer for α (any solution is valid as long as you are consistent, this is due to the difference between the solutions can be interpreted as a phase shift on the whole basis) we obtain,

$$\begin{aligned} \alpha &= \frac{\sqrt{1-\tan^2\theta}}{\sqrt{2}}, \\ \beta &= \frac{\sqrt{1-\tan^2\theta}}{\sqrt{2}}, \\ \gamma &= -\tan\theta. \end{aligned} \tag{4.22}$$

Now we have a unitary matrix U_3

$$U_3 = \begin{pmatrix} \frac{\tan\theta}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{\sqrt{1-\tan^2\theta}}{\sqrt{2}} \\ \frac{\tan\theta}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \frac{\sqrt{1-\tan^2\theta}}{\sqrt{2}} \\ \sqrt{1-\tan^2\theta} & 0 & -\tan\theta \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \tan\theta & 1 & \sqrt{1-\tan^2\theta} \\ \tan\theta & -1 & \sqrt{1-\tan^2\theta} \\ \sqrt{2(1-\tan^2\theta)} & 0 & -\sqrt{2}\tan\theta \end{pmatrix} \tag{4.23}$$

to be decomposed into beamsplitter like operations. The first T matrix will be T_{32}, but as we can see from U_3 , $U_{32} = 0$ so we don't need to apply anything to get it to zero. This means $T_{32} = I$ and therefore doesn't change U_3 at all as we require. The next reduction

will be on the U_{31} position as follows,

$$U_3' = U_3.T_{31} = \frac{1}{\sqrt{2}} \begin{pmatrix} \tan\theta & 1 & \sqrt{1-\tan^2\theta} \\ \tan\theta & -1 & \sqrt{1-\tan^2\theta} \\ \sqrt{2(1-\tan^2\theta)} & 0 & -\sqrt{2}\tan\theta \end{pmatrix} . \begin{pmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b^* & 0 & a^* \end{pmatrix}$$
$$= \frac{1}{\sqrt{2}} \begin{pmatrix} a\tan\theta + b^*\sqrt{1-\tan^2\theta} & 1 & -b\tan\theta + a^*\sqrt{1-\tan^2\theta} \\ a\tan\theta + b^*\sqrt{1-\tan^2\theta} & -1 & -b\tan\theta + a^*\sqrt{1-\tan^2\theta} \\ a\sqrt{2(1-\tan^2\theta)} - b^*\sqrt{2}\tan\theta & 0 & -b\sqrt{2(1-\tan^2\theta)} - a^*\sqrt{2}\tan\theta \end{pmatrix} . \tag{4.24}$$

To make the U_3' zero in the 3,1 position we require,

$$a\sqrt{2(1-\tan^2\theta)} - \sqrt{2}b^*\tan\theta = 0, \tag{4.25}$$

along with the condition $|a^2| + |b^2| = 1$ we have a simultaneous equation with a solution being,

$$a = \tan\theta \qquad \text{and} \qquad b = \sqrt{1-\tan^2\theta}. \tag{4.26}$$

Substituting these into (4.24) we get,

$$T_{31} = \begin{pmatrix} \tan\theta & 0 & -\sqrt{1-\tan^2\theta} \\ 0 & 1 & 0 \\ \sqrt{1-\tan^2\theta} & 0 & \tan\theta \end{pmatrix} \qquad \text{and} \qquad U_3' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} . \tag{4.27}$$

Repeating the process with T_{21} we have,

$$U_3'' = U_3'T_{21} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} . \begin{pmatrix} a & -b & 0 \\ b^* & a^* & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \frac{1}{\sqrt{2}} \begin{pmatrix} a+b^* & -b+a^* & 0 \\ a-b^* & -b-a^* & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} . \tag{4.28}$$

To satisfy $a - b^* = 0$ and $|a|^2 + |b|^2 = 1$ we can have the solution $a = b = 1/\sqrt{2}$ and

$$T_{21} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} , \tag{4.29}$$

leading to,

$$U_3'' = D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{4.30}$$

The two minus signs can be changed by a phase shift on the qubits at the end of the beamsplitter setup. If we use photons in spatial modes to represent each basis state, then the initial unitary U_3 can be implemented by a beamsplitter with reflection and transmission coefficients of $\tan\theta$ and $\sqrt{1-\tan^2\theta}$ respectively. These are between the first basis state and the auxiliary basis state, followed by a $50/50$ beamsplitter between the first and second basis.

4.4 Decomposing The Unambiguous Two Out Of Four Elimination Measurement

In section 3.3.3 we derived the measurement operators for the elimination of two out of the four two-qubit states, $|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta-,\theta\rangle$. At the limiting point when $\cos(2\theta) = \sqrt{2}-1$ the operators are proportional to the projectors onto the pure states,

$$
\begin{aligned}
|\psi_A\rangle &= \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}|00\rangle - \frac{1}{\sqrt{2}}|10\rangle,\\
|\psi_B\rangle &= \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}|00\rangle - \frac{1}{\sqrt{2}}|01\rangle,\\
|\psi_C\rangle &= \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}|00\rangle + \frac{1}{\sqrt{2}}|01\rangle,\\
|\psi_D\rangle &= \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}|00\rangle + \frac{1}{\sqrt{2}}|10\rangle,\\
|\psi_E\rangle &= \left(1-\frac{1}{\sqrt{2}}\right)|00\rangle + \frac{1}{\sqrt{2}}|11\rangle,\\
|\psi_F\rangle &= \left(1-\frac{1}{\sqrt{2}}\right)|00\rangle - \frac{1}{\sqrt{2}}|11\rangle,
\end{aligned}
\tag{4.31}
$$

which can be represented in matrix form as

$$
\begin{pmatrix}
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\\
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & -\frac{1}{\sqrt{2}} & 0 & 0\\
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{\sqrt{2}} & 0 & 0\\
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\
1-\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}}\\
1-\frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}}
\end{pmatrix},
\tag{4.32}
$$

where the rows make up the pure states $|\psi_A\rangle, ..., |\psi_F\rangle$, with the projector onto each state being proportional to the measurement operators $\Pi_A, ..., \Pi_F$. Each measurement operator eliminates the pair of states denoted by the label in it's subscript. These are,

$$
\begin{aligned}
A =&\{|\theta,\theta\rangle, |\theta,-\theta\rangle\},\\
B =&\{|\theta,\theta\rangle, |-\theta,\theta\rangle\},\\
C =&\{|\theta,-\theta\rangle, |-\theta,-\theta\rangle\},\\
D =&\{|-\theta,\theta\rangle, |-\theta,-\theta\rangle\},\\
E =&\{|\theta,-\theta\rangle, |-\theta,\theta\rangle\},\\
F =&\{|\theta,\theta\rangle, |-\theta,-\theta\rangle\}.
\end{aligned} \tag{4.33}
$$

The columns of (4.32) correspond to the basis states $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ from left to right. The first stage of the decomposition is to complete the Neumark extension by adding the two auxiliary basis states required to get a unitary matrix. If we extend by adding two the auxiliary basis states $|aux_1\rangle$ and $|aux_2\rangle$ then labelling each component of the basis states using a_i and b_i we can calculate the composition of this basis as shown by the following,

$$
\begin{pmatrix}
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 & a_1 & b_1\\
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & a_2 & b_2\\
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & a_3 & b_3\\
\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & a_4 & b_4\\
1-\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & a_5 & b_5\\
1-\frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & a_6 & b_6
\end{pmatrix}, \tag{4.34}
$$

giving us the requirements for orthogonality as,

$$
\begin{aligned}
&\left(\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}\right)(a_1+a_2+a_3+a_4)+(1-\frac{1}{\sqrt{2}})(a_5+a_6)=0,\\
&-\frac{a_2}{\sqrt{2}}+\frac{a_3}{\sqrt{2}}=0,\\
&-\frac{a_1}{\sqrt{2}}+\frac{a_4}{\sqrt{2}}=0,\\
&\frac{a_5}{\sqrt{2}}-\frac{a_6}{\sqrt{2}}=0,\\
&a_1^2+a_2^2+a_3^2+a_4^2+a_5^2+a_6^2=1.
\end{aligned} \tag{4.35}
$$

The same conditions apply for $|aux_2\rangle$ by replacing a_i with b_i, with the extra condition

$$
\sum_{i=1}^{6} a_i b_i = 0. \tag{4.36}
$$

Solving the middle three conditions of (4.35) we have

$$
a_2 = a_3, \qquad a_1 = a_4, \qquad a_5 = a_6, \tag{4.37}
$$

and substituting these in to the top and bottom conditions of (4.35) we get

$$2a_1\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}+2a_2\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}+2a_5(1-\frac{1}{\sqrt{2}})=0 \qquad \text{and} \qquad \sum_i a_i^2=1. \quad (4.38)$$

A simple solution is to make one of a_1, a_2 or a_5 zero, one option is then setting $a_1 = a_4 = -a_2 = -a_3$ and $a_5 = a_6 = 0$. This gives us,

$$a_1=\frac{1}{2},\ a_2=-\frac{1}{2},\ a_3=-\frac{1}{2},\ a_4=\frac{1}{2},\ a_5=0,\ a_6=0. \quad (4.39)$$

Here we have assumed the coefficients are real as a solution with real coefficients necessarily exists and so this assumption is made for ease of calculations. It also seems intuitive that complex solutions would not give more elements as zero in the decomposition. This is definitely something to look into further though. Now b_i must satisfy the conditions in (4.35) as well as $\sum_{i=1}^{6} a_i b_i = 0$. From these we can see

$$\begin{aligned} &b_2=b_3, \qquad b_1=b_4, \qquad b_5=b_6, \\ &\frac{1}{2}(b_1+b_4-b_2-b_3)=0. \end{aligned} \quad (4.40)$$

To satisfy these conditions we set $b_1 = b_2 = b_3 = b_4$. This leaves us with the following two simultaneous equations arising from the top and bottom conditions from (4.35)

$$4b_1\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}+2b_5(1-\frac{1}{\sqrt{2}})=0 \qquad \text{and} \qquad 4b_1^2+2b_5^2=1. \quad (4.41)$$

The two solutions are $b_1 = \pm(1/2 - 1/\sqrt{2})$ with $b_5 = \pm\sqrt{\sqrt{2}-1}$. Taking the positive solution gives us the unitary

$$U_6=\begin{pmatrix} \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ 1-\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & \sqrt{\sqrt{2}-1} \\ 1-\frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & \sqrt{\sqrt{2}-1} \end{pmatrix}. \quad (4.42)$$

A program in Matlab was written to decompose the matrix into T matrices. The aim was to have the minimum number of non-identity T matrices and as it is possible to permute the rows and columns and keep the function of the unitary the same, it is possible to reduce the number of non-identity T matrices. We go into more detail about optimising this process in chapter 4.6.

We tested a large collection of matrices we thought may give good results, and in the end

found the following unitary gave a simple decomposition,

$$U_{opt} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \sqrt{\sqrt{2}-1} & 1-\frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \sqrt{\sqrt{2}-1} & 1-\frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} \end{pmatrix}, \tag{4.43}$$

where the new rows and columns representing the pure states and the basis state are

$$\begin{pmatrix} |\psi_F\rangle \\ |\psi_B\rangle \\ |\psi_A\rangle \\ |\psi_C\rangle \\ |\psi_E\rangle \\ |\psi_D\rangle \end{pmatrix} \qquad \begin{pmatrix} |11\rangle & |01\rangle & |10\rangle & |aux_1\rangle & |aux_2\rangle & |00\rangle \end{pmatrix} \quad . \tag{4.44}$$

There are ten zeroes in under the diagonal of U_{opt} and in the decomposition process from Matlab these became identity matrices so we can already give the following ten T matrices as

$$T_{62} = T_{61} = T_{54} = T_{53} = T_{52} = T_{43} = T_{41} = T_{32} = T_{31} = T_{21} = I. \tag{4.45}$$

It is not a given that elements that begin as zero will be identity matrices in the final decomposition, as multiplication by the T_{ij} matrices affects the i and j columns, but it seems with our unitary this effect is nullified. The remaining five T matrices are

$$T_{65} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.91 & -0.414 \\ 0 & 0 & 0 & 0 & 0.414 & 0.91 \end{pmatrix},$$

$$T_{63} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \; T_{42} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$T_{64} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \; T_{51} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.46}$$

If we instead took the negative solutions for b_i then we would have the initial matrix given by

$$U_6 = \begin{pmatrix} \sqrt{\frac{1}{\sqrt{2}} - \frac{1}{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & -(\frac{1}{2} - \frac{1}{\sqrt{2}}) \\ \sqrt{\frac{1}{\sqrt{2}} - \frac{1}{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{2} & -(\frac{1}{2} - \frac{1}{\sqrt{2}}) \\ \sqrt{\frac{1}{\sqrt{2}} - \frac{1}{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & -(\frac{1}{2} - \frac{1}{\sqrt{2}}) \\ \sqrt{\frac{1}{\sqrt{2}} - \frac{1}{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{2} & -(\frac{1}{2} - \frac{1}{\sqrt{2}}) \\ 1 - \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & -\sqrt{\sqrt{2} - 1} \\ 1 - \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & -\sqrt{\sqrt{2} - 1} \end{pmatrix}. \tag{4.47}$$

After running the Matlab program the results were all the same with the exception of some global phase shifts on the beamsplitters. This is as expected, therefore we can take the positive solutions and confidently know this will not affect the final result.

4.5 Implementation

Our generic beamsplitter is represented by

$$U(2) = e^{i\phi} \begin{pmatrix} a & -b \\ b^* & a^* \end{pmatrix}, \tag{4.48}$$

where a and b are the transmission and reflection coefficients respectively. If we have our two incoming modes represented by,

$$|\alpha\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |\beta\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.49}$$

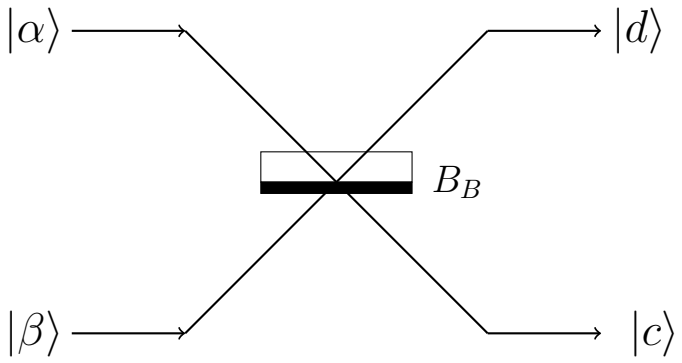


Figure 4.1: In this figure we have a beamsplitter with a phase shift upon reflection off the bottom. The output of the beamsplitter is described by equations (4.50) *and* (4.51).

for the top and bottom modes respectively. Applying a 50/50 beamsplitter, where $a = b = 1/\sqrt{2}$, a photon incident in the top mode gives us,

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\left[\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right] = \frac{1}{\sqrt{2}}(|c\rangle + |d\rangle), \tag{4.50}$$

then similarly for a photon incident in the lower mode we obtain,

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\left[\begin{pmatrix}-1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right] = \frac{1}{\sqrt{2}}(-|c\rangle + |d\rangle), \tag{4.51}$$

where c and d are the two output ports as shown in figure 4.1. In equation (4.51) one of the outcoming photons has picked up a phase shift of π so has a minus sign as the phase $e^{i\pi} = -1$. We make sure we label this shifted output as the reflected to account for the phase shift that occurs upon reflection by a half silvered mirror. This is then a $50/50$ beamsplitter with a phase shift upon reflection off the bottom that we shall denote by B_{50B}. Similarly a $50/50$ beamsplitter with a phase shift upon reflection off the top will be denoted by B_{50T} and these beamsplitters are given by

$$B_{50T} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix} \qquad \text{and} \qquad B_{50B} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}. \tag{4.52}$$

In our T matrices from (4.46) we have minus signs in many different places and this affects which way round the the mirrors will be and also whether certain inputs or outputs will require a phase shift as well. T_{65} is of the form of the beamsplitter described in (4.50), so this will be a beamsplitter with a phase shift upon the bottom, except with a transmission of $0.91^2 \approx 0.83$ and reflection of proportion $0.414^2 \approx 0.17$.
We then have the beamsplitter parts of T_{64}, T_{53} and T_{42} as,

$$T_{64} \equiv T_{53} \equiv T_{42} \equiv \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix} = B_{50T} \tag{4.53}$$

Then finally we have the relevant part of T_{51} as

$$T_{51} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} e^{i\pi} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = -B_{50B}. \tag{4.54}$$

As we can see this is just B_{50B} with a phase shift, therefore we can just apply a B_{50B} and put two π phase shifters on the input or output arms. A linear optical setup could look something like figure 4.2.

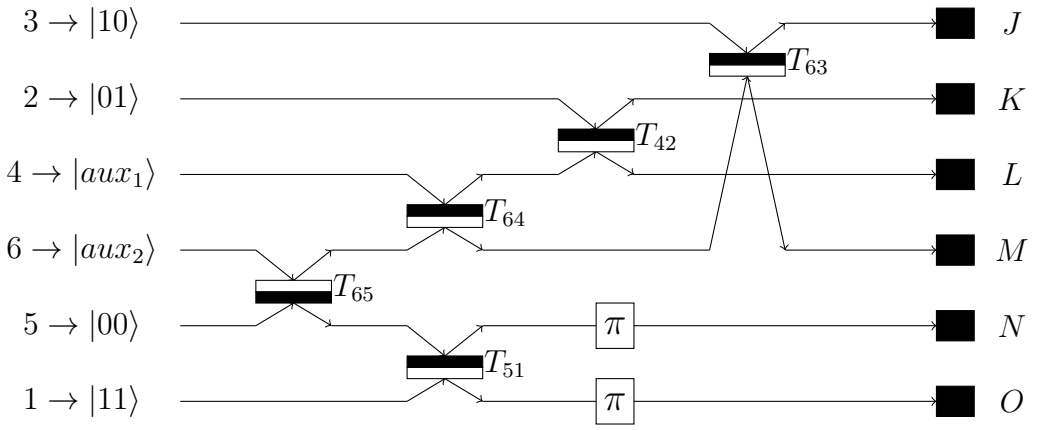


Figure 4.2: The setup of the T_{ij} matrices from (4.46)*, with the input modes on the left and detectors at the output on the right. Reflection upon the black side of the mirror introduces a phase shift of π. All the matrices excluding T_{65} are 50/50 beamsplitters. The ordering of the input states has been done for clarity to minimise the number of lines crossing each other. A click at each detector represents the elimination of one of the possible pairings from* (4.33)*.*

In this case the modes 5 and 1 don't interact with anything again after T_{51} so the phase shifts are not required and in this scenario the same result can be been achieved with,

$$T_{51} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.55}$$

This is something that should try to be initially but our fundamental aim was to minimise the number of T matrices and currently there isn't a cost to having

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} \qquad \text{instead of} \qquad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{4.56}$$

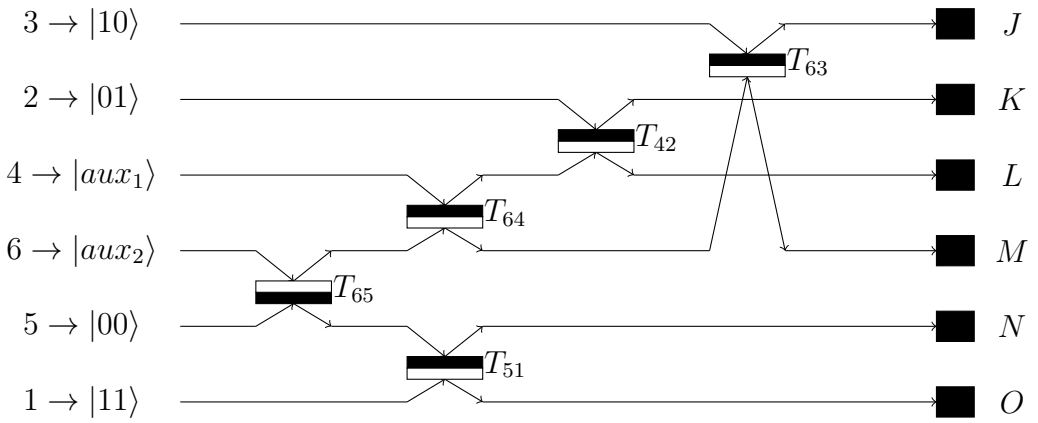


Figure 4.3: The improved implementation from Figure 4.2 after the removal of the unnecessary phase shifters.

Figure 4.3 shows a simple implementation when using spatial modes and linear optics.

Encoding each tensor product state as a single spatial mode removes the tensor product structure of the original problem but in this case we only wish to eliminate sets involving the product state so are not interested in specifically observing the first or second qubit. If we wished to do this we would have to find a different implementation.

To see which output relates to which measurement we can follow the basis states through the system, monitoring the phase shifts, then we can which pure state each outcome relates to by the sign of each basis state. This allows us a quick method for finding which output corresponds to which elimination result. For example if we look at the J outcome we can see the inputs that can make it to J are $|10\rangle, |aux_1\rangle, |00\rangle$ and $|aux_2\rangle$. As the auxiliary states are vacuum the final outcomes do not depend on these. So if we follow the $|10\rangle$ path to J it picks up a negative phase shift from T_{63} so we have $-|10\rangle$. Then for $|00\rangle$ it does not pick up any negative phase shifts from T_{65} , T_{64} and T_{63}, giving us $|00\rangle$. The output J has components of $|00\rangle$ and $-|01\rangle$, this is then proportional to $|\psi_A\rangle$ from equation (4.31). For the six outputs labelled $J - O$ in figure 4.3 we have,

$$\begin{aligned}
&J \to |00\rangle - |10\rangle \propto |\psi_A\rangle && \text{Eliminates} && \{|\theta,\theta\rangle, |\theta,-\theta\rangle\},\\
&K \to |00\rangle - |01\rangle \propto |\psi_B\rangle && \text{Eliminates} && \{|\theta,\theta\rangle, |-\theta,\theta\rangle\},\\
&L \to |00\rangle + |01\rangle \propto |\psi_C\rangle && \text{Eliminates} && \{|\theta,-\theta\rangle, |-\theta,-\theta\rangle\},\\
&M \to |00\rangle + |10\rangle \propto |\psi_D\rangle && \text{Eliminates} && \{-|\theta,\theta\rangle, |-\theta,-\theta\rangle\},\\
&N \to -|00\rangle - |11\rangle \propto |\psi_E\rangle && \text{Eliminates} && \{|\theta,-\theta\rangle, |-\theta,\theta\rangle\},\\
&O \to -|00\rangle + |11\rangle \propto |\psi_F\rangle && \text{Eliminates} && \{|\theta,\theta\rangle, |-\theta,-\theta\rangle\},
\end{aligned} \tag{4.57}$$

where $|\psi_i\rangle$ are the states from (4.31) and therefore if we get a click at the detector in position J we can say we have eliminated the states related to the measurement operator

Π_A which is $\{|\theta,\theta\rangle, |\theta,-\theta\rangle\}$, and similarly for the other detectors.

It also fairly simple to check the success probabilities for each outcome in the implementation. Equation (3.85) show the probabilities of each outcome should be $p(A,B,C,D) = (1/2)\cos(2\theta) \approx 0.21$ and $p(E,F) = 1/2 - \cos(2\theta) \approx 0.08$.
The four states $|\theta,\theta\rangle, |\theta,-\theta\rangle, |-\theta,\theta\rangle, |-\theta-\theta\rangle$ written in terms of $|00\rangle, |10\rangle, |01\rangle$ and $|11\rangle$ are

$$\begin{aligned}
|\theta,\theta\rangle &= \cos^2\theta|00\rangle + \cos\theta\sin\theta|01\rangle + \cos\theta\sin\theta|10\rangle + \sin^2\theta|11\rangle,\\
|\theta,-\theta\rangle &= \cos^2\theta|00\rangle - \cos\theta\sin\theta|01\rangle + \cos\theta\sin\theta|10\rangle - \sin^2\theta|11\rangle,\\
|-\theta,\theta\rangle &= \cos^2\theta|00\rangle + \cos\theta\sin\theta|01\rangle - \cos\theta\sin\theta|10\rangle - \sin^2\theta|11\rangle,\\
|-\theta,-\theta\rangle &= \cos^2\theta|00\rangle - \cos\theta\sin\theta|01\rangle - \cos\theta\sin\theta|10\rangle + \sin^2\theta|11\rangle. \qquad (4.58)
\end{aligned}$$

The states differ due to the signs but the magnitudes of the state in each basis state is the same. These being $|\cos^2\theta|$ for $|00\rangle$, $|\cos\theta\sin\theta|$ for $|01\rangle$ and $10\rangle$ and finally $|\sin^2\theta|$ for $|11\rangle$. The square of these magnitudes can be used to check the probabilities of each measurement outcome. As we will have a single photon state we can calculate the probability of each outcome. To calculate the probabilities we do a similar process to how we calculated which states were eliminated. For example if we take outcome J then we know we just have to consider the $|00\rangle$ and $|01\rangle$ inputs. We can see that probability of outcome J calculated from the proportions that reach J for our angle $\theta = 32.76°$ is

$$p(J) = \cos^4\theta(0.83 \times 0.5 \times 0.5) + \cos^2\theta\sin^2\theta(0.5) \approx 0.21. \qquad (4.59)$$

The 0.83 occurs from the transmission of T_{65} and the two following 0.5s occur from T_{64} and T_{63}. The 0.5 for the $\cos^2\theta\sin^2\theta$ term comes from T_{63}. This is the same result from our calculation in (3.85) and therefore is a good check our implementation works as planned. Looking at figure 4.3 we can see outcomes J, K, L and M all have $|00\rangle$ transmitted with probability 0.83 and then two 50/50 beamsplitters, accompanied by one of $|01\rangle$ or $|10\rangle$ going through a single 50/50 beamsplitter. Therefore all these will give the same probability as J as expected. For N and O we have

$$p(N) = p(M) = \cos^4\theta(0.17 \times 0.5) + \sin^4\theta(0.5) \approx 0.08. \qquad (4.60)$$

The sum of the probabilities add up to one as required. This method of calculating the probabilities is valid in this scenario as it is equivalent to calculating the conditional probability of each outcome depending on a single photon in an input mode. Then scaling these with the probabilities of each input.

We can also see the probability of each outcome given each input state

	J	K	L	M	N	O
$\|00\rangle$	0.83/4	0.83/4	0.83/4	0.83/4	0.17/2	0.17/2
$\|01\rangle$	0	0.5	0	0.5	0	0
$\|10\rangle$	0.5	0	0.5	0	0	0
$\|11\rangle$	0	0	0	0	0.5	0.5

Table 4.1: The probabilities of each outcome $J - O$ given on the four input states $|00\rangle, |01\rangle, |10\rangle, |11\rangle$.

In this implementation we have the six dimensions required all as spatial modes. It is of course possible to use other possibilities, another method using linear optics is to combine spatial modes and polarisation simultaneously. This can allow you to have two modes in the same path, each with different polarisations.

4.5.1 State Preparation

In the previous section we formulated an experimental set-up capable of implementing the measurement, but we also require a method to prepare the initial state and this is what I will go through below. The aim is to produce the states $|\theta, \theta\rangle, |\theta, -\theta\rangle, |-\theta, \theta\rangle, |-\theta-, \theta\rangle$, which are written out in equation (4.58). To form these we can take four spatial modes with each one representing $|00\rangle, |01\rangle, |10\rangle$ or $|11\rangle$ and splitting up a single input into the correct magnitudes for each arm, with a phase shifter at the end of all the arms except $|00\rangle$. By turning the phase shifters on or off when required we can produce any of the four states. This is shown in figure 4.4.

As we need to split the initial mode into four spatial modes ($|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$), where the modes $|00\rangle$ and $|01\rangle$ have the same amplitude. It seems sensible that the first beamsplitter of figure 4.4 should separate the $|01\rangle$ and $|10\rangle$ from $|00\rangle$ and $|11\rangle$. For the layout in figure 4.4 this would require a reflection proportion of $2\cos\theta\sin\theta$ and transmission $\cos^4\theta + \sin^4\theta$. Then $|01\rangle$ and $|10\rangle$ need to split evenly so BS_3 is simply a $50/50$ beamsplitter. Then for BS_2 we need to split it with a ratio of $\cos^2\theta/\sin^2\theta$ so we chose to transmit the $\cos^2$ proportion. The $1/\sqrt{\cos^4\theta + \sin^4\theta}$ is required to make sure the beam-

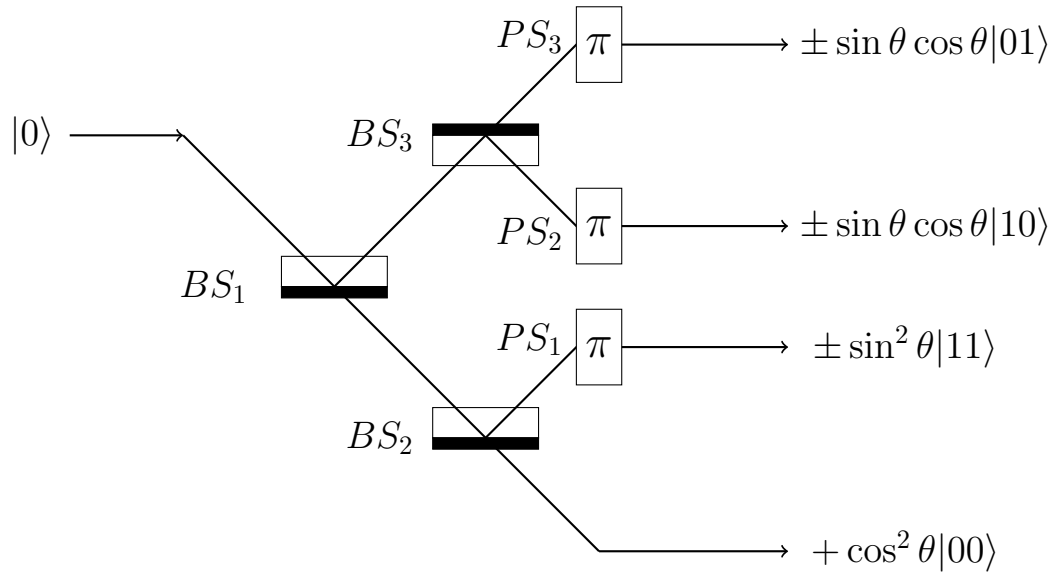


Figure 4.4: The state preparation system starts with an input from the left and then the beamsplitters as given in equation (4.61) *form the proportion of the state in each basis* $|00\rangle, |01\rangle, |10\rangle$ *and* $|11\rangle$ *with amplitudes* $\cos^2\theta, \cos\theta\sin\theta, \cos\theta\sin\theta$ *and* $\sin^2\theta$ *respectively. Reflection upon the black side introduces a phase shift, yet in this set-up there will be no phase shifts due to reflection. The phase shifters introduce a phase shift of* π *that introduces the minus sign. These will be turned on or off depending on which state we wish to produce.*

splitter is a unitary operator. Using this the three beamsplitters are,

$$
\begin{aligned}
BS_1 &= \begin{pmatrix} \sqrt{\cos^4\theta + \sin^4\theta} & -\sqrt{2}\cos\theta\sin\theta \\ \sqrt{2}\cos\theta\sin\theta & \sqrt{\cos^4\theta + \sin^4\theta} \end{pmatrix}, \\
BS_2 &= \frac{1}{\sqrt{\cos^4\theta + \sin^4\theta}} \begin{pmatrix} \cos^2\theta & -\sin^2\theta \\ \sin^2\theta & \cos^2\theta \end{pmatrix}, \\
BS_3 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.
\end{aligned} \tag{4.61}
$$

The transmission and reflection coefficients can be read off from the general matrix,

$$
BS = \begin{pmatrix} t & -r \\ r* & t* \end{pmatrix} \qquad \text{where} |r|^2 + |t|^2 = 1. \tag{4.62}
$$

$|r|^2$ and $|t|^2$ are the proportion of light that gets reflected or transmitted, or alternatively the probability a photon will get reflected or transmitted.

The table below shows how to produce each state with the phase shifters on or off.

State	PS_1	PS_2	PS_3
$\|\theta,\theta\rangle$	Off	Off	Off
$\|\theta,-\theta\rangle$	On	Off	On
$-\|\theta,\theta\rangle$	On	On	Off
$\|-\theta,-\theta\rangle$	Off	On	On

Then from this preparation the output arms need to match the input arms from figure 4.3 and then depending on which detector clicks you will eliminate one of the six possible pairs of states.
There will be other methods to prepare the state but this setup seemed to be an intuitive approach to preparing it in a simple manner.

4.5.2 Variable Beamsplitter

$50/50$ beamsplitters are fairly commonplace in most linear optics laboratories yet in both our measurement and state preparation we require beamsplitters with specific reflection and transmission coefficients different to that of a $50/50$ beamsplitter. It is possible to get these produced by a company or to buy a tunable beamsplitter, but often it can be easier or cheaper to use devices you already have. Figure (4.5) shows how to make a variable beamsplitter out of two $50/50$ beamsplitters, two mirrors and a variable phase shifter. This is called a Mach-Zehnder interferometer and can also be used to determine the phase shift caused by some sample by measuring the outputs. The degree of the phase shift is altered by changing the width or substance of the phase shifter.
If we have our input states

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.63}$$

then after the first beamsplitter we get

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{4.64}$$

The phase shifter can be represented as

$$PS = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}. \tag{4.65}$$

As the mirrors give an equal phase shift to both arms this can be interpreted as a global phase shift and does not need to be accounted for. Therefore the state before the second beamsplitter is

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\phi} \end{pmatrix}, \tag{4.66}$$

then after the final beamsplitter we get

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\phi} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 - e^{i\phi} \\ 1 + e^{i\phi} \end{pmatrix}. \tag{4.67}$$

Comparing this to the output of a generic beamsplitter we have

$$\begin{pmatrix} a & -b \\ b^* & a^* \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b^* \end{pmatrix}, \tag{4.68}$$

therefore

$$a = \frac{1}{2}(1 - e^{i\phi}) \qquad \text{and} \qquad b^* = \frac{1}{2}(1 - e^{i\phi}), \tag{4.69}$$

where a and b are the transmission and reflection coefficients respectively. If we represent the whole system by U_{vb} then we get

$$\begin{aligned} U_{vb}|0\rangle &= \frac{1}{2}[(1 - e^{i\phi})|0\rangle + (1 + e^{i\phi})|1\rangle], \\ U_{vb}|1\rangle &= -\frac{1}{2}[(1 + e^{i\phi})|0\rangle + (1 - e^{i\phi})|1\rangle], \end{aligned} \tag{4.70}$$

where the inputs and outputs are shown in figure (4.5)

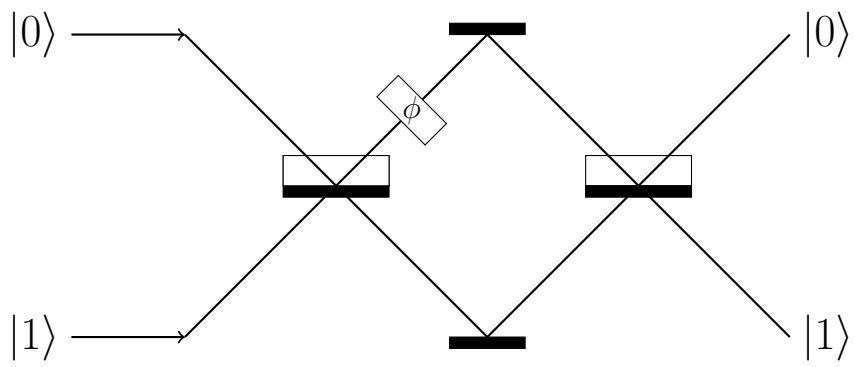


Figure 4.5: The setup for producing a variable beamsplitter using $50/50$ beamsplitters, mirrors and a phase shifter

Another possibility is to utilise polarisation to implement a variable beamsplitter. This can be implemented by a wave plate followed by a polarising beamsplitter (PBS). The transmission and reflection coefficients are dependent on the rotation of the waveplate. After the PBS the polarisations of each path will be different. These can of course be rotated back to the desired polarisation or if it doesn't matter they can be just left as they are. This is in fact what Thorlabs [59] sell as a tunable variable beamsplitter.

4.6 Optimising the Decomposition

4.6.1 What Is Optimal?

Whilst finding the decomposition for the unambiguous two out of four measurement we thought about how to optimise the decomposition process for any measurement. The first

stage is to consider what the desired setup is for a measurement, i.e how we define a better or worse setup. Our primary aim is to reduce the number of optical elements required to implement the measurement. This means maximising the amount of the T matrices that are the identity matrix or a global phase shift of the identity matrix. Furthermore we know that $50/50$ beamsplitters are cheaper and more readily available than ones with a variable reflectivity. We have seen variable beam splitters can be implemented using $50/50$ beamsplitters and phase shifters, or polarising beam splitters and a waveplate, but in each of these cases it requires multiple elements. Therefore a system with more $50/50$ beamsplitters will be deemed better.

Fewer elements is optimal obviously from a cost perspective but also more significantly from a performance aspect. A standard non-polarising 50/50 beamsplitter can have an error of $\pm 10\%$ in the reflection and transmission coefficients [59]. These errors can propagate throughout the system and sometimes certain elements in a system are more susceptible to causing a larger reduction in the fidelity of an experiment [60]. Other implementation approaches are of course possible including integrated optics, or using spin systems and this approach could potentially be extended to those methods with a small alteration of the optimisation.

In the reduction process we aim to make the off diagonal elements 0. In the method I described this was done by fitting the beamsplitter values so that each element under the diagonal became zero. The maximum number of elements equal to zero under the diagonal in an NxN matrix is $N(N-1)/2$. As we have choice in our auxiliary basis states and can permute the rows and columns of our unitary (as shown in the next section), we can attempt to have as many elements equal zero as possible under the diagonal of the unitary. In the next section we shall go about showing how to implement this strategy as well as looking into the relationship between the number of zeroes under the diagonal and the number of optical elements required.

4.6.2 How To Optimise

The starting point is the M X N matrix from (4.2) with the M rows consisting of the pure states that correspond to the rank-one measurement operators in the N different dimensions. This matrix contains all the information required to make the measurement. From here the next step is to apply the Neumark extension if required. If $N < M - 1$ then we will have choice for at least one of the basis states. After this we also have the ability to permute the columns and rows of the unitary as long as the content of each row and column remains the same. This is just the equivalent to a relabelling of the states. For

example we can take the 4 X 4 matrix,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \tag{4.71}$$

then permute the rows to give us,

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{pmatrix}, \tag{4.72}$$

followed by permuting the columns to give us,

$$\begin{pmatrix} a_{23} & a_{21} & a_{22} & a_{24} \\ a_{43} & a_{41} & a_{42} & a_{44} \\ a_{33} & a_{31} & a_{32} & a_{34} \\ a_{13} & a_{11} & a_{12} & a_{14} \end{pmatrix}. \tag{4.73}$$

Our approach is to maximise the amount of zeroes in the auxiliary basis states. This has not been proven to be the best way to have as many elements equal zero as possible under the diagonal but it seems to be a logical approach. After multiplication by each T_{ij} matrix the number of zeroes is not necessarily preserved. It is possible that elements that are originally zero can become non-zero when a different element is made to be zero.

In the two out of four case the number of zeroes regularly correlated with the number of identity T matrices. Yet when you multiply by T_{ij} you alter the whole of the i and j columns with each multiplication. In the two out of four case we noticed a correlation between number of zeroes under the diagonal and the number of non-identity T_{ij} matrices so it seemed to give a good result but not always the best. Finding a relationship between the unitary matrix and the number of non-identity T_{ij} matrices is definitely a next step in this procedure otherwise it may be a requirement to perform the decomposition on each permutation that would significantly increase the computing resources required.

4.6.3 Method Of Optimisation

The optimisation process was written into a Matlab code by Ittoop Puthoor and myself with the idea that you could input your M X N matrix and then as an output you would receive the collection of T matrices that form the best experimental setup. So far it is at the stage where you can input an M X M matrix and with minimal adaptation of the code required, the specific T matrices are outputted. The obvious next step is to have the

program perform the Neumark extension and optimise U alongside the row and column permutations.

The code is split up into different parts. These are creating the general T matrices, sorting the M X M matrix U_M to give as many zeroes as possible under the diagonal and finally decomposing the sorted matrix to find the specific T matrices. The code for the first two parts were predominantly completed by Ittoop with the decomposition being equally worked upon.

To create the general T matrices a set of M X M identity matrices are formed and then filled in with a and b components from (4.3) in the correct positions dependent on the values of i and j to create the T_{ij} matrices like in (4.4).

The matrix sort begins with comparing the number of zeroes in the top row of U_M with the bottom row. If there are more zeroes in the top row then it is swapped with the bottom row and if there are less zeroes in the top row then it is left the same. This then repeated with the second row and the bottom, the third row and the bottom, all the way up until the $M-1$ row is being compared with the bottom row. At the end of this procedure the row with the most zeroes in will be on the bottom row. The bottom row is then fixed and then the procedure is repeated comparing the first row and the $M-1$ row so that the row with the second most zeroes is in the second from bottom row. You repeat this procedure $M-1$ times so that all the rows are now ordered from least zeroes to most zeroes from top to bottom respectively. This is done as the bottom row has a larger contribution to the number of elements under the diagonal.
The next step is to perform the same procedure but comparing the columns and so we aim to get more zeroes to the left hand column as that is the one that contributes the most elements to underneath the diagonal. This seems to be an efficient and effective sort but we have managed to formulate special cases where the optimal solution sneaks through. Another option is to run through every permutation and store the number of zeroes under the diagonal along with the respective matrix and then after each permutation you compare and if the new matrix has more zeroes under the diagonal you then store the new one. There are $M!^2$ permutations for each unitary. Then a quick check is performed to verify that the final product is a unitary matrix as required.

Once we have the sorted U_M we decompose it into the T matrices. This is done by setting up the requirement that the targeted position in the matrix becomes zero and that $|a^2| + |b^2| = 1$ as is in done for USD in equation (4.25) for the position $3, 1$. These equations are then solved by Matlab and the T_{ij} matrix is produced with solutions for a and b. A new unitary matrix is then formed after by multiplying the initial unitary by T_{ij}

and then the procedure is repeated until we have completely decomposed U_M into the T_{ij} matrices and D.

There are still further ways to both optimise this process and make it easier for the user. Mostly finding an explicit relationship between the unitary matrix and the number of non-identity T matrices in the final decomposition. Also specifying the costs (both monetary and error inducing) of using certain optical elements and minimising this factor in the decomposition as well as just the number of non-identity matrices.

We did apply the optimisation method to the two out of four elimination measurement and as an outcome got

$$\begin{pmatrix} 0 & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{2} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{2} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{2} & \sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 1-\frac{1}{\sqrt{2}} & \sqrt{\sqrt{2}-1} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 1-\frac{1}{\sqrt{2}} & \sqrt{\sqrt{2}-1} \end{pmatrix}, \tag{4.74}$$

After decomposing this we got six non-identity T matrices. This is still a fairly good result but not as simple as the one we found using trial and error. This shows there is still work to be done but in general whilst trying out scenarios there was a trend between more zeroes under the diagonal and the number of identity matrices in the decomposition. It wasn't a guaranteed method for a good result though.

4.7 Conclusion

In this chapter we have introduced the method of decomposing a unitary matrix into beamsplitter like operations so that an experimental set up can be produced for any measurement. With this we produced a fairly simple looking approach to implementing our unambiguous two out of four elimination method and now we are starting to approach experimentalists to see whether implementation is feasible.
Furthermore we looked at optimising the decomposition process for any measurement so that a program could be run to find the simplest or at least a simple implementation method. This work has potential but so far there isn't any definite answers as how best to optimise it.

Chapter 5

Joint Measurements

A joint measurement is when a single measurement on a single quantum system gives a result for each of two observables. A measurement of two observables is when the outcome consists of results for both of the observables. In a simple case a measurement on one of the observables is performed and then the result of the other observable is guessed at random. This may not be the optimal joint measurement but it gives you a result for each of the two observables by performing a single measurement. If the two observables commute with each other then a standard Von-Neumann projective measurement [5] will accomplish the task involved. With this in mind the interesting case to consider is joint measurements on non-commuting observables.

The sharpness of a joint measurement determines how well you measure a single observable. For example if we performed a projective Von-Neumann measurement on one observable this would have maximum sharpness of one, whilst if we made a random guess of the other observable this would have a minimum sharpness of zero. There are different ways to determine an optimal measurement with different trade-off relationships between the sharpness on each observable.

The motivation for our work came from a collaboration with an experimental group in Bristol led by Adetunmise Dada. The aim of the experiment was to experimentally test a trade-off relationship for the sharpness of qubit measurements. An area of error for the experiment was that when trying to achieve a certain sharpness for each measurement the actual value differed due to experimental error. The experimentalists were just measuring an eigenstate of one of the observables each time. We were tasked with investigating what is the best state to measure to reduce the error in the estimates of the sharpness. This was called the probe state so we set about finding the optimal probe state.

The experiment was performed with a relatively simple optical implementation without the requirement of filtering, post-selection or entanglement with an ancilla. It used a

high quality heralded single photon source and was able to produce results extremely similar to the quantum mechanical limit of how much the variance must increase when performing joint measurements.

Our aim was to find the optimal probe state that would minimise the uncertainty in our estimation of the sharpness of a measurement between two observables. This work was done by myself, Ittoop Puthoor and Erika Andersson with the experimental work led by Adetunmise Dada in Bristol [61].

5.1 BB84 Example

In section 3.4.2 we introduced the BB84 quantum key distribution protocol. In this section we briefly look at what happens if the eavesdropper Eve attempted to measure the states in either the $\{|0\rangle, |1\rangle\}$ basis, which we shall call the vertical basis, or the $\{|-\rangle, |+\rangle\}$ basis, which we will call the diagonal basis. If we just look for Eve to attempt to learn the state and not worry about her cheating being detected, then when Alice announces the basis she sent the states in, Eve's measurement results would be equivalent to a random guess when she measured in the wrong basis and a correct result when she measured in the correct basis. The probability of success for this method with each basis being chosen with equal probability by Alice is 0.75, as 50% of the time she will measure in the correct basis, gaining the correct result and when she measures in the wrong basis there is a 50% chance of her guessing the correct state.

Another method is performing a different joint measurement on the BB84 states. For equal a priori probabilities of each basis an obvious choice for a single basis measurement would be a basis halfway between $\{|0\rangle, |1\rangle\}$ and $\{|-\rangle, |+\rangle\}$. This basis is given by the basis states

$$\begin{aligned} |a\rangle &= \frac{\sqrt{2-\sqrt{2}}}{2}|0\rangle + \frac{\sqrt{2+\sqrt{2}}}{2}|1\rangle, \\ |b\rangle &= \frac{\sqrt{2+\sqrt{2}}}{2}|0\rangle - \frac{\sqrt{2-\sqrt{2}}}{2}|1\rangle. \end{aligned} \tag{5.1}$$

In the $\{|a\rangle, |b\rangle\}$ basis we would attribute an outcome related to $|a\rangle$ with the outcome $|+\rangle$ in the diagonal basis and $|1\rangle$ in the vertical basis. Similarly an outcome related to $|b\rangle$ would be attributed to $|-\rangle$ in the diagonal basis and $|0\rangle$ in the vertical basis. The probabilities of each result are given by the square of the overlaps between the states that compose the

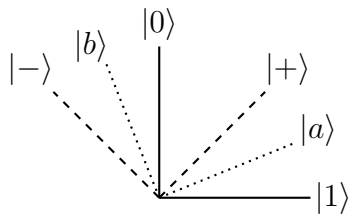


Figure 5.1: The four BB84 states $|0\rangle, |1\rangle, |+\rangle$ and $|-\rangle$ with the new measurement basis $|a\rangle$ and $|b\rangle$ as defined in equation (5.1).

basis. These probabilities are,

$$
\begin{aligned}
|\langle 0|a\rangle|^2 &= \frac{2-\sqrt{2}}{4} \approx 0.146 \qquad \text{and} \qquad |\langle 0|b\rangle|^2 = \frac{2+\sqrt{2}}{4} \approx 0.854, \\
|\langle 1|a\rangle|^2 &= \frac{2+\sqrt{2}}{4} \approx 0.854 \qquad \text{and} \qquad |\langle 1|b\rangle|^2 = \frac{2-\sqrt{2}}{4} \approx 0.146, \\
|\langle +|a\rangle|^2 &= \left| \frac{1}{\sqrt{2}} \left(\frac{\sqrt{2-\sqrt{2}}}{2} + \frac{\sqrt{2+\sqrt{2}}}{2} \right) \right|^2 \approx 0.854, \\
|\langle +|b\rangle|^2 &= \left| \frac{1}{\sqrt{2}} \left(\frac{\sqrt{2+\sqrt{2}}}{2} - \frac{\sqrt{2+\sqrt{2}}}{2} \right) \right|^2 \approx 0.146, \\
|\langle -|a\rangle|^2 &= \left| \frac{1}{\sqrt{2}} \left(\frac{\sqrt{2-\sqrt{2}}}{2} - \frac{\sqrt{2+\sqrt{2}}}{2} \right) \right|^2 \approx 0.146, \\
|\langle -|b\rangle|^2 &= \left| \frac{1}{\sqrt{2}} \left(\frac{\sqrt{2-\sqrt{2}}}{2} + \frac{\sqrt{2+\sqrt{2}}}{2} \right) \right|^2 \approx 0.854.
\end{aligned} \tag{5.2}
$$

As expected $|\langle 0|a\rangle|^2 + |\langle 0|b\rangle|^2 = 1$ and similarly for the other probabilities giving a total probability of one for each scenario.

Assuming Alice picks her basis at random, this would give an average probability of picking the correct bit as 85.4%. This is an immediate improvement upon picking a basis and guessing. It is possible there is a more optimal joint measurement to perform, it seems for measuring in a single basis this is most likely to be best. As an attempt to eavesdrop in QKD it is flawed by the fact that Eve is altering the states during every measurement so when Alice and Bob declare some of their bits it will be found the states have been tampered with.

This joint measurement is not too dissimilar to quantum random access codes (QRAC) [62, 63, 64] in which the aim is to encode n classical bits into m qubits and the receiver aims to recover one of the bits. This is given by the notation $n \xrightarrow{p} m$, where $p > 1/2$ is the probability of success. In the above measurement we have two outcomes encoded into one measurement and so is similar to $2 \xrightarrow{p} 1$, where $p = 0.85$. In fact this is the maximum success probability for a $2 \xrightarrow{p} 1$ QRAC.

5.2 Spin 1/2 Observables

For a spin $1/2$ system we have the observables $\hat{A} = \mathbf{a} \cdot \hat{\sigma}$ and $\hat{B} = \mathbf{b} \cdot \hat{\sigma}$, where **a** and **b** are unit vectors on the Bloch sphere. The observables have eigenvalues of $+1$ and -1 which will be assigned to the results spin up and spin down respectively. Often a condition imposed is that the expectation values of the jointly measured observables are proportional to the expectation values of the observables as if they were measured separately. This is called the joint unbiasedness condition and is employed in other joint measurement work [65][66]. This gives us,

$$\langle \hat{A}_J \rangle = \alpha \langle \hat{A} \rangle \qquad \text{and} \qquad \langle \hat{B}_J \rangle = \beta \langle \hat{B} \rangle, \tag{5.3}$$

where the subscript J refers to the joint measurement and it must hold for the coefficients that $0 < |\alpha|, |\beta| < 1$. Ideally we want α and β to be as close to one as possible as this makes the expectation values from the jointly measured observables as close as possible to those from the observables measured separately.
Of course α and β can't both be one for non-commuting observables otherwise you would be able to distinguish between non-orthogonal quantum states. The condition relating the sharpness of the measurement to the observables derived by Busch and Lahti [67], and also derived more generally by Andersson et al. [68] is,

$$|\alpha\mathbf{a} + \beta\mathbf{b}| + |\alpha\mathbf{a} - \beta\mathbf{b}| \leq 2. \tag{5.4}$$

A geometrical interpretation of the bound is given by the authors in [68], and is shown in figure 5.2, where the sum of the lengths of the diagonals must be less than two. Given **a** and **b** are unit vectors, this also leads to the requirements that $|\alpha|$ and $|\beta|$ must be less than one. This can easily be seen as vector addition gives

$$|\alpha\mathbf{a} + \beta\mathbf{b}| = d_2 \qquad \text{and} \qquad |\alpha\mathbf{a} - \beta\mathbf{b}| = d_1, \tag{5.5}$$

where d_1 and d_2 are shown in figure 5.2.

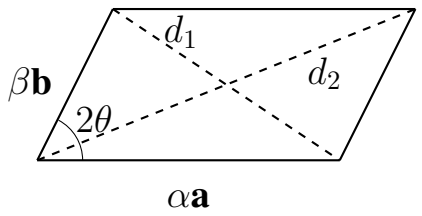


Figure 5.2: The bound given in (5.4) *equates to the sum of the diagonals in this parallelogram with sides of length $\alpha\mathbf{a}$ and $\beta\mathbf{b}$ being less than or equal to length 2. α and β can both equal one only when **a** and **b** are parallel and the parallelogram just becomes a straight line and therefore the sum of the diagonals is twice the length of that straight line.*

Furthermore by squaring equation (5.4) twice Andersson et al. [68] showed the bound can be written 5.4 as,

$$\Delta_\alpha^2 \Delta_\beta^2 = \frac{(1-\alpha^2)(1-\beta^2)}{\alpha^2\beta^2} \leq \sin^2(2\theta), \tag{5.6}$$

where,

$$\cos 2\theta = \mathbf{a}\cdot\mathbf{b} \qquad , \qquad \Delta_\alpha^2 \equiv \frac{1-\alpha^2}{\alpha^2} \qquad \text{and} \qquad \Delta_\beta^2 \equiv \frac{1-\beta^2}{\beta^2}. \tag{5.7}$$

The bound in (5.6) is an uncertainty relation where the uncertainty is purely from the fact that $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are quantum observables measured jointly. This is similar to the Heisenberg-Schrödinger-Robertson uncertainty except in contrast this bound is tight and is independent of the measured state. It is only dependent on the measured quantum observables and only valid for spin 1/2 particles. The bound in equation (5.4) assumes the joint measurement has marginal measurement operators of the form,

$$\Pi_\pm^a = \frac{1}{2}(\hat{\mathbf{1}} \pm \alpha\mathbf{a}\cdot\hat{\sigma}) \qquad \text{and} \qquad \Pi_\pm^b = \frac{1}{2}(\hat{\mathbf{1}} \pm \beta\mathbf{b}\cdot\hat{\sigma}), \tag{5.8}$$

as is done in [67]. A joint measurement with the marginal measurement operators given in (5.8) with any α and β that saturate the bound in equation (5.4) can always be realised.

5.3 Joint Measurement of a Spin 1/2 System

Andersson et al. [68] proposed measuring the two spin 1/2 systems $\hat{A}$ and $\hat{B}$ by performing a projective measurement along one of two directions **c** or **d** as shown in figure 5.3 with respective probabilities p and $1-p$. If the measurement outcomes are ± 1 for each observable, then there are four possible outcomes,

$$\{A=1, B=1\}\ ,\ \{A=-1, B=1\}\ ,\ \{A=1, B=-1\} \text{ or } \{A=-1, B=-1\}. \tag{5.9}$$

Similarly the outcome of a measurement in the **c** or **d** direction will give outcomes of $C=\pm 1$ or $D=\pm 1$ respectively. Table 5.1 shows how we associate the results of C or D with those of A and B.

The expectation values for the joint measurement are therefore,

$$\begin{aligned} \bar{A}_J &= p\langle \mathbf{c}\cdot\hat{\sigma}\rangle + (1-p)\langle \mathbf{d}\cdot\hat{\sigma}\rangle, \\ \bar{B}_J &= p\langle \mathbf{c}\cdot\hat{\sigma}\rangle - (1-p)\langle \mathbf{d}\cdot\hat{\sigma}\rangle. \end{aligned} \tag{5.10}$$

If we use the marginal measurement operators from equation (5.8) instead, then we require

$$\bar{A}_J = \alpha\langle \mathbf{a}\cdot\hat{\sigma}\rangle \qquad \text{and} \qquad \bar{B}_J = \beta\langle \mathbf{b}\cdot\hat{\sigma}\rangle. \tag{5.11}$$

C or D	A	B
C=+1	+1	+1
C=-1	-1	-1
D=+1	+1	-1
D=-1	-1	+1

Table 5.1: Table showing the relationship between the outcomes of the measurements on C or D and the results for A and B.

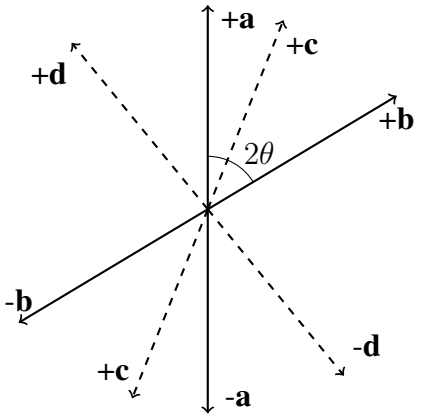


*Figure 5.3: The two spin observables $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ have direction **a** and **b** on the Bloch sphere, where a result in the positive direction equates to +1 and -1 in the negative direction. **c** and **d** represent the directions used in the joint measurement of $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$. The angle 2θ is that between **a** and **b**.*

For these two expectation values to be equivalent, we can obtain a relationship between **a,b** and **c,d**. Equating the expectation values from (5.10) and (5.11) gives us,

$$\begin{aligned}\alpha\langle\mathbf{a}\cdot\hat{\sigma}\rangle &= p\langle\mathbf{c}\cdot\hat{\sigma}\rangle + (1-p)\langle\mathbf{d}\cdot\hat{\sigma}\rangle,\\ \beta\langle\mathbf{b}\cdot\hat{\sigma}\rangle &= p\langle\mathbf{c}\cdot\hat{\sigma}\rangle - (1-p)\langle\mathbf{d}\cdot\hat{\sigma}\rangle.\end{aligned} \tag{5.12}$$

Solving this for **d** we get,

$$\begin{aligned}p\langle\mathbf{c}\cdot\hat{\sigma}\rangle &= \beta\langle\mathbf{b}\cdot\hat{\sigma}\rangle + (1-p)\langle\mathbf{d}\cdot\hat{\sigma}\rangle,\\ \alpha\langle\mathbf{a}\cdot\hat{\sigma}\rangle &= \beta\langle\mathbf{b}\cdot\hat{\sigma}\rangle + 2(1-p)\langle\mathbf{d}\cdot\hat{\sigma}\rangle,\\ \langle\mathbf{d}\cdot\hat{\sigma}\rangle &= \frac{\alpha\langle\mathbf{a}\cdot\hat{\sigma}\rangle - \beta\langle\mathbf{b}\cdot\hat{\sigma}\rangle}{2(1-p)},\end{aligned} \tag{5.13}$$

and similarly for **c** we get

$$\langle\mathbf{c}\cdot\hat{\sigma}\rangle = \frac{\alpha\langle\mathbf{a}\cdot\hat{\sigma}\rangle + \beta\langle\mathbf{b}\cdot\hat{\sigma}\rangle}{2p}. \tag{5.14}$$

As both are being taken as an average over products with $\hat{\sigma}$ we can simplify to,

$$\mathbf{c} = \frac{(\alpha\mathbf{a}+\beta\mathbf{b})}{2p} \qquad \text{and} \qquad \mathbf{d} = \frac{(\alpha\mathbf{a}-\beta\mathbf{b})}{2(1-p)}. \tag{5.15}$$

As **c** and **d** are unit vectors that means $|\mathbf{c}| = |\mathbf{d}| = 1$, therefore

$$\frac{1}{2p}|\alpha\mathbf{a} + \beta\mathbf{b}| = 1 \qquad \text{and} \qquad \frac{1}{2(1-p)}|\alpha\mathbf{a} - \beta\mathbf{b}| = 1, \tag{5.16}$$

giving us,

$$p = \frac{|\alpha\mathbf{a} + \beta\mathbf{b}|}{2} \qquad \text{and} \qquad 1 - p = \frac{|\alpha\mathbf{a} - \beta\mathbf{b}|}{2}. \tag{5.17}$$

If we add these two equations together then we get,

$$1 = \frac{|\alpha\mathbf{a} + \beta\mathbf{b}| + |\alpha\mathbf{a} - \beta\mathbf{b}|}{2},$$
$$|\alpha\mathbf{a} + \beta\mathbf{b}| + |\alpha\mathbf{a} - \beta\mathbf{b}| = 2. \tag{5.18}$$

This is now a saturation of the bound given in equation (5.4), therefore showing a joint measurement performed by measuring the observables $\hat{C} = \mathbf{c} \cdot \hat{\sigma}$ or $\hat{D} = \mathbf{d} \cdot \hat{\sigma}$ with probabilities p and $1 - p$ respectively is optimal.

The measurements along **c** and **d** with probability p and $1 - p$ can be described by the measurement operators

$$\Pi^c_{\pm} = \frac{p}{2}(\hat{\mathbf{1}} \pm \mathbf{c}.\hat{\sigma}) \qquad \text{and} \qquad \Pi^c_{\pm} = \frac{1-p}{2}(\hat{\mathbf{1}} \pm \mathbf{d}.\hat{\sigma}). \tag{5.19}$$

Using these measurement operators, and the outcomes described in equations (5.1) and (5.17) that give p and $1 - p$ in terms of α, β, **a** and **b**, we can give measurement operators for the four possible outcomes associated with $\hat{A}$ and $\hat{B}$ as

$$\Pi^{ab}_{+1,+1} = \frac{1}{4}|\alpha\mathbf{a} + \beta\mathbf{b}|\hat{\mathbf{1}} + \frac{1}{4}|\alpha\mathbf{a} + \beta\mathbf{b}|.\hat{\sigma},$$
$$\Pi^{ab}_{+1,-1} = \frac{1}{4}|\alpha\mathbf{a} - \beta\mathbf{b}|\hat{\mathbf{1}} + \frac{1}{4}|\alpha\mathbf{a} - \beta\mathbf{b}|.\hat{\sigma},$$
$$\Pi^{ab}_{-1,+1} = \frac{1}{4}|\alpha\mathbf{a} - \beta\mathbf{b}|\hat{\mathbf{1}} - \frac{1}{4}|\alpha\mathbf{a} - \beta\mathbf{b}|.\hat{\sigma},$$
$$\Pi^{ab}_{-1,-1} = \frac{1}{4}|\alpha\mathbf{a} + \beta\mathbf{b}|\hat{\mathbf{1}} - \frac{1}{4}|\alpha\mathbf{a} + \beta\mathbf{b}|.\hat{\sigma}. \tag{5.20}$$

The next step is to calculate the sharpness for each observable finding α and β that saturate the bound in (5.6) in terms of p and the angle θ that is shown in figure 5.3. This can be done by using the equations from (5.17), and by minimising the error from α and β by making the bound in (5.6) an equality. As the magnitude of a vector can be given as,

$$|\mathbf{a} + \mathbf{b}| = \sqrt{|\mathbf{a} + \mathbf{b}|^2}, \tag{5.21}$$

we can then state,

$$|\alpha\mathbf{a} \pm \beta\mathbf{b}| = \sqrt{\alpha^2|\mathbf{a}|^2 + \beta^2|\mathbf{b}|^2 \pm 2\alpha\beta\mathbf{a}.\mathbf{b}}\,, \tag{5.22}$$

giving us,

$$p = \frac{1}{2}\sqrt{\alpha^2|\mathbf{a}|^2 + \beta^2|\mathbf{b}|^2 + 2\alpha\beta\mathbf{a}.\mathbf{b}}\,,$$
$$1 - p = \frac{1}{2}\sqrt{\alpha^2\mathbf{a}^2 + \beta^2\mathbf{b}^2 - 2\alpha\beta\mathbf{a}.\mathbf{b}}\,. \tag{5.23}$$

If we then square both of these terms to eliminate the square roots we get,

$$p^2 = \frac{1}{4}(\alpha^2|\mathbf{a}|^2 + \beta^2|\mathbf{b}|^2 + 2\alpha\beta\mathbf{a.b}) \tag{5.24}$$

$$1 + p^2 - 2p = \frac{1}{4}(\alpha^2|\mathbf{a}|^2 + \beta^2|\mathbf{b}|^2 - 2\alpha\beta\mathbf{a.b})\,. \tag{5.25}$$

By substituting (5.24) into (5.25) we get,

$$1 - 2p + \frac{1}{4}(\alpha^2|\mathbf{a}|^2 + \beta^2|\mathbf{b}|^2 + 2\alpha\beta\mathbf{a.b}) = \frac{1}{4}(\alpha^2|\mathbf{a}|^2 + \beta^2|\mathbf{b}|^2 - 2\alpha\beta\mathbf{a.b})\,,$$

$$2p - 1 = \alpha\beta\mathbf{a.b}. \tag{5.26}$$

We can relate the vectors **a** and **b** by the variable θ defining their separation as $\mathbf{a.b} = |\mathbf{a}||\mathbf{b}|\cos(2\theta)$. If we take **a** and **b** as unit vectors then $|\mathbf{a}| = |\mathbf{b}| = 1$ giving us,

$$2p - 1 = \alpha\beta\cos(2\theta). \tag{5.27}$$

Therefore

$$\alpha = \frac{2p-1}{\beta\cos(2\theta)} \qquad \text{and} \qquad \beta = \frac{2p-1}{\alpha\cos(2\theta)}, \tag{5.28}$$

and from equation (5.6) we know the optimal solution will be given when

$$\frac{(1-\alpha^2)(1-\beta^2)}{\alpha^2\beta^2} = \left(\frac{1}{\alpha^2} - 1\right)\left(\frac{1}{\beta^2} - 1\right) = \sin^2(2\theta), \tag{5.29}$$

Substituting in the term for α from (5.28) we get,

$$\left(\frac{\beta^2\cos^2(2\theta)}{(2p-1)^2} - 1\right)\left(\frac{1}{\beta^2} - 1\right) = \sin^2(2\theta),$$

$$\frac{\cos^2(2\theta)}{(2p-1)^2} - \frac{\beta^2\cos^2(2\theta)}{(2p-1)^2} - \frac{1}{\beta^2} + 1 = \sin^2(2\theta),$$

$$\frac{\beta^2\cos^2(2\theta)}{(2p-1)^2} - \frac{\beta^4\cos^2(2\theta)}{(2p-1)^2} - 1 + \beta^2 = \beta^2\sin^2(2\theta),$$

$$\beta^4 - \beta^2\left[1 + \frac{(2p-1)^2}{\cos^2(2\theta)} - \sin^2(2\theta)\frac{(2p-1)^2}{\cos^2(2\theta)}\right] + \frac{(2p-1)^2}{\cos^2(2\theta)} = 0. \tag{5.30}$$

As $1 - \sin^2(2\theta) = \cos^2(2\theta)$, the equation immediately above can be simplified to

$$\beta^4 - \beta^2[1 + (2p-1)^2] + \frac{(2p-1)^2}{\cos^2(2\theta)} = 0. \tag{5.31}$$

We can solve for β^2 using the quadratic formula and then taking the square root again we obtain the value for β as

$$\beta_{opt} = \pm\sqrt{\pm\sqrt{[2(p-1)p+1]^2 - (1-2p)^2\sec^2(2\theta)} + 2(p-1)p + 1}, \tag{5.32}$$

and then for α we have,

$$\alpha_{opt} = \frac{(2p-1)}{\beta\cos(2\theta)}, \tag{5.33}$$

where the subscript opt has been adopted to clarify that these are α and β solutions that saturate the bound given in (5.6). There are four possible solutions for both α_{opt} and β_{opt} depending on the signs chosen for in front of the square roots. We will refer to these solutions as $++, +-, -+$ and $--$ for the chosen signs respectively. We shall just consider the two positive solutions $++$ and $+-$ as the negative solutions have the same magnitude as one of the positive solutions. α_{opt} and β_{opt} will saturate the bound given in (5.4). Looking at the parallelogram in figure 5.2 we can also see that to saturate the bound we are choosing **c** and **d** such that the diagonals are $2p$**c** and $2(1-p)$**d** and this will maximise α and β. This is because using equations (5.15) and (5.17) we can write

$$2p\mathbf{c} = (\alpha\mathbf{a} + \beta\mathbf{b}) \qquad \text{and} \qquad 2(1-p)\mathbf{d} = (\alpha\mathbf{a} - \beta\mathbf{b}). \tag{5.34}$$

If we look at a couple of cases for specific p values we can see what measurement this performs and the values for α_{opt} and β_{opt}. For $p = 1$ we measure only in the **c** direction and the magnitude of β_{opt} for any solution starts at $|\beta_{opt}| = 1$ for $\theta = 0$ but then increases. As $|\beta_{opt}| \leq 1$ the only time we can satisfy the bound in equation (5.4) for $p = 1$ is when $\theta = 0$ and **a=b**. Here we see that there are limitations on our value of θ from fixing p. In the $p = 1$ case it would be obvious to measure along the direction of **a** or **b**. It also shows that only measuring along one of **c** and **d** can't produce the optimal joint measurement for the cases when $\theta > 0$. As we will show in the following chapter the bounds on α and β are not the only factors limiting the angle θ.

For $p = 1/2$ the solution for β_{opt} will always be undefined as the term $(1-2p)^2 \sec^2(2\theta)$ is equivalent to $0/0$ so there is no valid solution. This is because if we look at the fact that the length of the diagonals must be $2p$**c** and $2(1-p)$**d** then we can see the diagonals must have equal length and this can only be satisfied by a rectangle. Therefore the angle between **a** and **b** must be $90°$ so $2\theta = 90°$. As $\sec(2\theta) = 1/\cos(2\theta) = 1/0$ for $2\theta = 90°$ and $(1-2p) = 0$ for $p = 1/2$ then we have this $0/0$ term appearing. Finally we will look at the case when $p = 0.7$ as this is the scenario that was used in the setup for the experimental realisation in [61]. Figure 5.4 shows the relationship between the magnitudes of α_{opt} and β_{opt}. As opposed to the $p = 1$ case, we have a large selection of angles θ for which we have valid α and β values to satisfy the bound from (5.4). Yet in the next chapter we shall see that for θ larger than at the point the values of α_{opt} and β_{opt} meet the measurement is not physical.

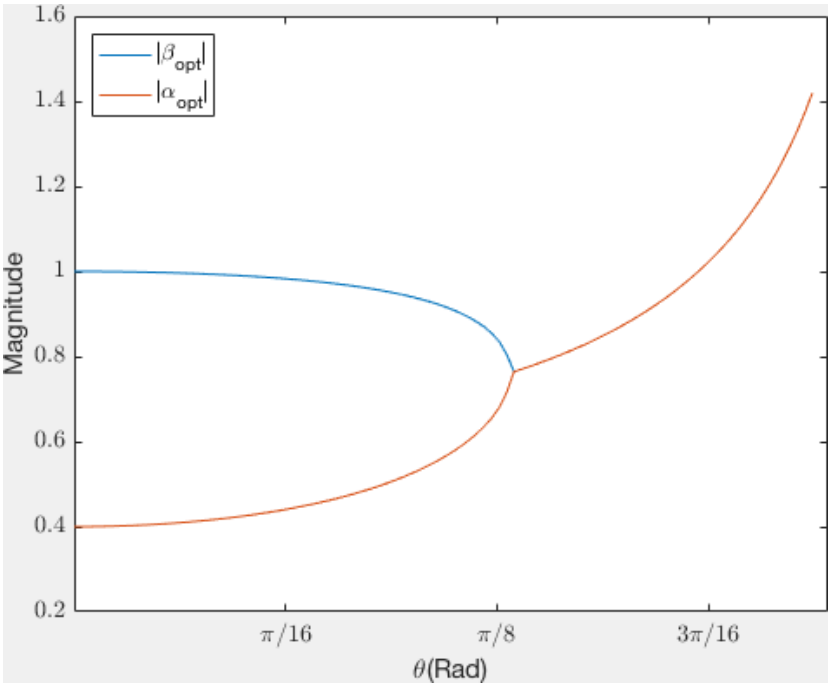


Figure 5.4: The magnitudes of β_{opt} and α_{opt} from (5.32) for the $++$ solution are plotted against θ for $p = 0.7$ in the range $0 \leq 2\theta \leq \pi/2$. The magnitudes for the $+-$ solutions are equivalent with the α and β terms switching.

5.3.1 Fixing the probability p

α_{opt} and β_{opt} are both dependent on the probability p that gives the probability whether we measure in the **c** direction (p) or in the **d** direction ($1-p$). The initial approach would be to vary the angle between the states to be measured and see how this affects the other variables. In the experimental setup with Dada et al. [61] the probability p was fixed and this then puts limits on the measurements that can be performed. So in our work we looked at what varies when we fix p to a certain value. From figure 5.4 we see that for large θ the magnitude of α and β exceed one so we know there are not always solutions that can saturate the bound in equation (5.4) for all values of θ. We shall also see that the requirement that **c** and **d** can't be separated by more than $90°$ puts limits on the physical values of θ.

If for example we set $p = 1/2$, which would be not just feasible, but probably the easiest probability split, as it would involve using a simple $50/50$ beamsplitter. The parallelogram in figure 5.2 will become a rectangle as the diagonals will both have length 1 as $|2p\mathbf{c}| = |2(1-p)\mathbf{d}| = 1$. This means the angle between **a** and **b** will have to be $90°$, and therefore can only represent an optimal joint measurement between two maximally complementary spin 1/2 observables. With $p = 1/2$ we can also realise any optimal α and β values with a suitably chosen the angle between **c** and **d**. If **c** and **d** are

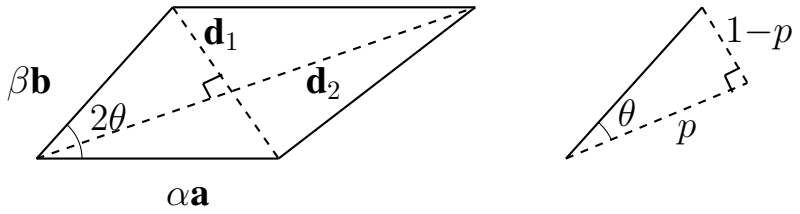


Figure 5.5: On the left we have the full parallelogram when the angle between **c** *and* **d** *is* 90° *leading to the maximum achievable angle* 2θ *between* **a** *and* **b**. $\mathbf{d}_1 = 2(1-p)\mathbf{c}$ *and* $\mathbf{d}_2 = 2p\mathbf{d}$. *On the right we just have the relevant right angled triangle required to calculate the maximum angle with the lengths of the sides.*

in a similar direction then the parallelogram approaches a straight line as one of α or β approaches a magnitude of one whilst the other becomes very small. This goes to the extent of when **c**=**d** then we are just measuring one observable sharply with $\alpha = 0$ and $\beta = 1$ or vice versa. If the sharpness is zero, this is equivalent to a random guess of the respective observable. In fact anything that satisfies $|\alpha|^2 + |\beta|^2 = 1$ will be an optimal joint measurement.

If $p > 1/2$ then the two diagonals of the parallelogram will have different lengths with the **c** direction being longer. Therefore the angle between **a** and **b** will necessarily have to be less than 90°. The maximum achievable angle will be achieved when the angle between **c** and **d** is 90°. This is the greatest the angle between **c** and **d** can be and still have a parallelogram. Figure 5.5 shows the maximum achievable angle 2θ can take. From simple trigonometry we can tell that

$$\tan(\theta_{max}) = \frac{1-p}{p}, \tag{5.35}$$

therefore the angle 2θ is restricted by the condition

$$2\theta \leq 2\arctan\left(\frac{1-p}{p}\right). \tag{5.36}$$

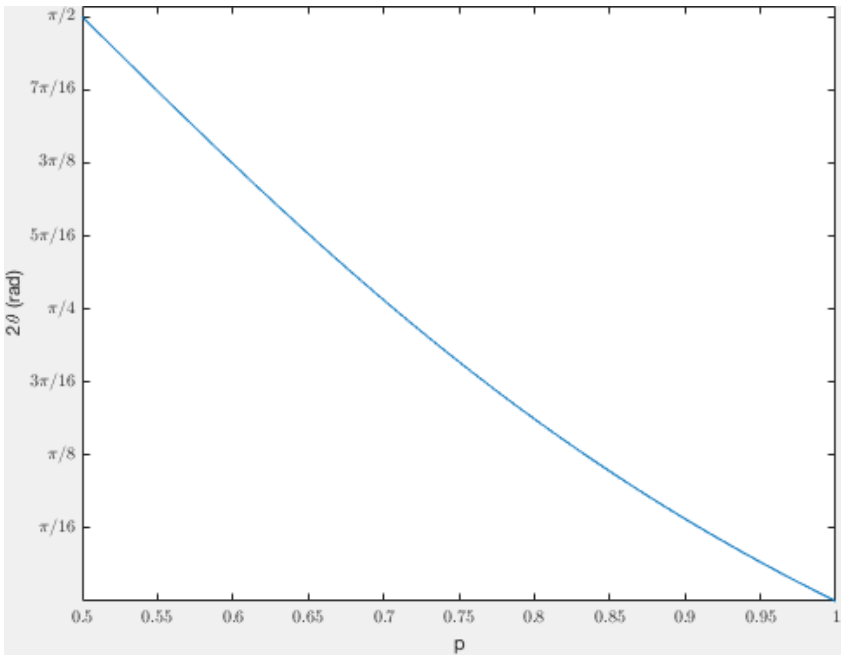


Figure 5.6: This figure shows how the maximum angle varies depending on our choice of p. We have chosen $p \geq 1/2$, this can be done without any loss of generality. As we have stated for $p = 1/2$ we have $2\theta = \pi/2$, and for $p = 1$ we have $2\theta = 0$.

In the next section when we look at optimising the probe state for specific p values.

5.4 Optimum Probe State

An aim of the experimental work by Dada et al.[61] was to estimate what the sharpness of a joint measurement is, from measurement results in an experiment. To do this they calculated **c** and **d** from the values of α_{opt} and β_{opt}. Due to experimental errors the sharpness couldn't be calculated exactly and in fact the values α_{exp} and β_{exp} were used. The difference between the optimal and experimental values differed depending on the measured (or probe) state.
Our aim was to find the optimum probe state that would give us the minimum error in the estimation of the sharpness. First of all we have to derive an expression for error in the sharpness of the measurement and then from that look at minimising that function.

We will start by calculating the variance in the estimate of $\langle \hat{A} \rangle$ of a single unsharp measurement and from that proceed to finding the errors and errors in the sharpness of a joint measurement. The definition of variance is given as

$$\Delta^2(x) = \sum_{i=1}^{M} p_i (x_i - \mu)^2, \tag{5.37}$$

where M is the possible number of outcomes, x_i is the value of each result and μ is the mean of the results. For a single (non-joint) measurement $\mu = \langle \hat{A} \rangle$ and $x_i = a_i$.

Substituting this in and expanding out the squared term we get

$$\Delta^2(\langle\hat{A}\rangle) = \left(\sum_{i=1}^{M} p_i a_i^2 - 2\sum_{i=1}^{M} p_i a_i \langle\hat{A}\rangle + \sum_{i=1}^{M} p_i \langle\hat{A}\rangle^2\right), \tag{5.38}$$

as $a_i = \pm 1$, then $\sum_i^M p_i a_i^2 = 1$. Also $\sum_i^M p_i a_i = \langle\hat{A}\rangle$ and so the second term becomes $-2M\langle\hat{A}\rangle^2$. This gives us the variance for measuring a single system as

$$\begin{aligned}\Delta^2(\langle\hat{A}\rangle) &= \left(1 - 2\langle\hat{A}\rangle^2 + \langle\hat{A}\rangle^2\right) \\ &= 1 - \langle\hat{A}\rangle^2.\end{aligned} \tag{5.39}$$

If we now measure N systems then the error reduces by a factor of $1/\sqrt{N}$ and if we take the error squared it is simply

$$\Delta_e^2(\langle\hat{A}\rangle) = \frac{\Delta^2(\langle\hat{A}\rangle)}{N} = \frac{1}{N}(1 - \langle\hat{A}\rangle^2), \tag{5.40}$$

where Δ_e^2 is used to denote the squared error of the variable in the parentheses and Δ^2 is used to denote the variance. Looking more closely at the squared error in 5.40 we can see if we have a system where $\langle\hat{A}\rangle = 1$ then our error will be zero. This is because for $\langle\hat{A}\rangle$ to equal one the system must be in spin-up every single time, therefore each time we measure this single observable sharply we will get $+1$ every time. The same occurs for $\langle\hat{A}\rangle = -1$ but with the system in spin down each time. If $\langle\hat{A}\rangle = 0$ then for each measurement there is a $50/50$ chance of obtaining either -1 or $+1$, therefore if we have a finite measurement there is a chance of an error. In the simplest case of $N = 1$ there will be a guaranteed difference of 1 between your measured value and the average. This also helps to see the relevance of the $\frac{1}{N}$ term, as the larger N is the lower the error will be.
However for a joint measurement, when we don't measure each observable sharply, we have a new average $\alpha\langle\hat{A}\rangle$ and the variance is

$$\Delta^2(\alpha\langle\hat{A}\rangle) = 1 - \alpha^2\langle\hat{A}\rangle, \tag{5.41}$$

and then measuring N systems the squared error is given by

$$\Delta_e^2(\alpha\langle\hat{A}\rangle) = \frac{1}{N}(1 - \alpha^2\langle\hat{A}\rangle^2). \tag{5.42}$$

As we are trying to estimate the sharpness of a measurement from an experiment we know $\langle\hat{A}\rangle$ but are now trying to find the error in α. As $\alpha = \alpha\langle\hat{A}\rangle/\langle\hat{A}\rangle$ then we can find the error by dividing by $\langle\hat{A}\rangle^2$ to give

$$\Delta_e^2(\alpha) = \frac{\Delta_e^2(\alpha\langle\hat{A}\rangle)}{\langle\hat{A}\rangle^2} = \frac{1}{N}\frac{(1 - \alpha^2\langle\hat{A}\rangle^2)}{\langle\hat{A}\rangle^2} = \frac{1}{N}\left(\frac{1}{\langle\hat{A}\rangle^2} - \alpha^2\right),$$

and similarly for error in $\hat{B}$, we have

$$\Delta_e^2(\beta) = \frac{\Delta_e^2(\beta\langle\hat{B}\rangle)}{\langle\hat{B}\rangle^2} = \frac{1}{N}\frac{(1 - \beta^2\langle\hat{B}\rangle^2)}{\langle\hat{B}\rangle^2} = \frac{1}{N}\left(\frac{1}{\langle\hat{B}\rangle^2} - \beta^2\right). \tag{5.43}$$

To minimise $\Delta_e^2(\alpha)$ or $\Delta_e^2(\beta)$ we want to set $\alpha = \pm 1$ or $\beta = \pm 1$ respectively, and this is equivalent to using eigenstates of $\hat{A}$ or $\hat{B}$ respectively. Yet often there will have to be a trade-off between the two errors and to minimise this error we will need to introduce a probe state. We shall call this probe state **p** and figure 5.7 shows the Bloch vector of the probe state and it's relations to **a** and **b** are

$$\mathbf{a.p} = |\mathbf{a}||\mathbf{p}| \cos\phi \qquad \text{and} \qquad \mathbf{b.p} = |\mathbf{b}||\mathbf{p}| \cos(2\theta - \phi). \tag{5.44}$$

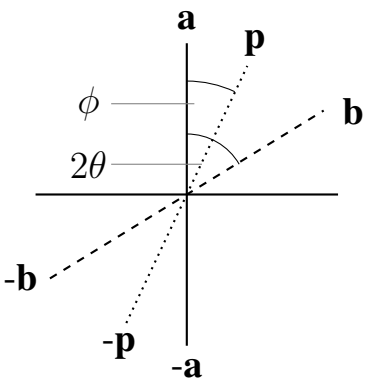


Figure 5.7: The two directions **a** *(solid line) and* **b** *(dashed line) are associated with the observables $\hat{A}$ and $\hat{B}$ respectively, with the separation angle of 2θ between them. Also* **p** *(dotted line) is the probe state and has a separation angle of ϕ from* **a**.

Now we have the individual errors that we wish to minimise but as it is a joint measurement we need to find a way of combining the two errors to create a joint error to minimise.

5.4.1 Sum Of Errors

One way to combine the errors is simply to sum them and then minimise

$$\Delta_e^2(\alpha) + \Delta_e^2(\beta) = \frac{1}{N}\left[\left(\frac{1}{\langle\hat{A}\rangle^2} - \alpha^2\right) + \left(\frac{1}{\langle\hat{B}\rangle^2} - \beta^2\right)\right]. \tag{5.45}$$

As N, α and β are fixed then we don't need to look at terms that have no $\langle\hat{A}\rangle$ or $\langle\hat{B}\rangle$ dependence, from this we can simplify the function we require to minimise to

$$e_{sum} = \frac{1}{\langle\hat{A}\rangle^2} + \frac{1}{\langle\hat{B}\rangle^2}. \tag{5.46}$$

Intuitively due to the symmetry of the function we want to minimise, it seems sending the probe state halfway between **a** and **b** would be optimal. We can check this and also run through the process of optimising the error that we will use in future cases.
From figure 5.7 we can see we have the following relations

$$\langle\hat{A}\rangle^2 = |\mathbf{a.p}|^2 = \cos^2(\phi) \qquad \text{and} \qquad \langle\hat{B}\rangle^2 = |\mathbf{b.p}|^2 = \cos^2(2\theta - \phi). \tag{5.47}$$

This gives us the sum of errors in terms of the variable ϕ as

$$\Delta_e^2(\alpha) + \Delta_e^2(\beta) = \frac{1}{N}\left[\left(\frac{1}{\cos^2(\phi)} - \alpha^2\right) + \left(\frac{1}{\cos^2(2\theta - \phi)} - \beta^2\right)\right]. \tag{5.48}$$

First if we use the eigenstate of $\hat{A}$ as the probe state, from figure 5.7 we see this gives us $\phi = 0$ and therefore $\mathbf{p} = \mathbf{a}$. In this scenario $\langle\hat{A}\rangle^2 = 1$ and $\langle\hat{B}\rangle^2 = \cos^2(2\theta)$. This will give the sum of errors as,

$$\Delta_e^2(\alpha) + \Delta_e^2(\beta) = \frac{1}{N}\left[\left(1 - \alpha^2\right) + \left(\frac{1}{\cos^2(2\theta)} - \beta^2\right)\right], \tag{5.49}$$

and similarly if we make $\mathbf{b} = \mathbf{p}$ by making $\phi = 2\theta$ then the sum of the errors is,

$$\Delta_e^2(\alpha) + \Delta_e^2(\beta) = \frac{1}{N}\left[\left(\frac{1}{\cos^2(2\theta)} - \alpha^2\right) + \left(1 - \beta^2\right)\right]. \tag{5.50}$$

Another scenario we can check is what happens if we send the state down the middle, which is equivalent to setting $\phi = \theta$. This will give us a sum of errors of,

$$\Delta_e^2(\alpha) + \Delta_e^2(\beta) = \frac{1}{N}\left[\left(\frac{1}{\cos^2\theta} - \alpha^2\right) + \left(\frac{1}{\cos^2\theta} - \beta^2\right)\right]. \tag{5.51}$$

Comparing sending the eigenstates of $\hat{A}$ or $\hat{B}$ or down the middle respectively, the errors

$$e_{(\phi=0,2\theta)} = 1 + \frac{1}{\cos^2(2\theta)} \qquad \text{and} \qquad e_{(\phi=\theta)} = \frac{2}{\cos^2\theta}. \tag{5.52}$$

Figure 5.8 shows us that the error when $\phi = \theta$ is lower than that when $\phi = 0$ or $\phi = 2\theta$ therefore sending the probe state between down the middle is better than sending one of the eigenstates. Even though we have shown sending the probe state down the middle of the observables is better than the eigenvalues, we have not proven it is optimal. One way to calculate this is to differentiate the reduced sum of the errors given in (5.46) with respect to the angle of the probe state ϕ. Then equating this to zero and checking it is a minimum we can find the value of ϕ that gives us the the smallest error for each value of θ. The minimisation is then

$$\begin{aligned} &\frac{\partial}{\partial\phi}\left(\frac{1}{\cos^2(\phi)} + \frac{1}{\cos^2(2\theta - \phi)}\right) = 0 \\ &= \frac{2\tan(\phi)}{\cos^2\phi} - \frac{2\tan(2\theta - \phi)}{\cos^2(2\theta - \phi)} = 0, \end{aligned} \tag{5.53}$$

and from this we can see that setting $\phi = \theta$ reaches the minimum, therefore sending the probe state directly in between **a** and **b** is optimal for minimising the sum of the errors given in (5.45).

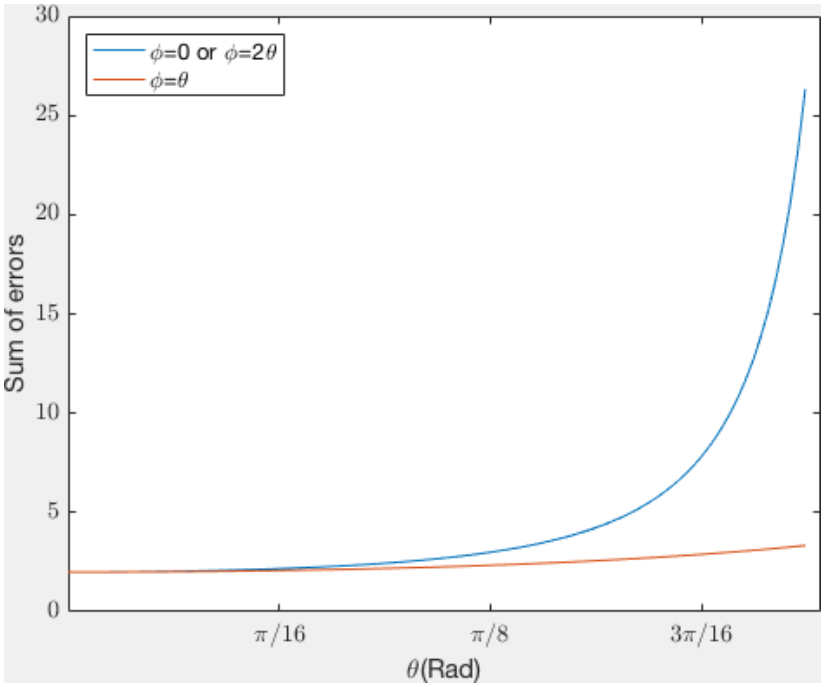


Figure 5.8: Comparison of the two errors when the probe state is either the eigenvalue of $\hat{A}$ or $\hat{B}$ or down the middle. The errors are in the reduced form given in (5.52)*. The error from the probe state is lower when sent down the middle than for the eigenvalues of the measured observables.*

5.4.2 Product Of Errors

Another possibility is to minimise the product of the errors given by,

$$\begin{aligned}\Delta_e^2(\alpha)\Delta_e^2(\beta) &= \frac{1}{N^2}\left[\left(\frac{1}{\langle\hat{A}\rangle^2}-\alpha^2\right)\left(\frac{1}{\langle\hat{B}\rangle^2}-\beta^2\right)\right], \\ &= \frac{1}{N^2}\left[\frac{1}{\langle\hat{A}\rangle^2\langle\hat{B}\rangle^2}-\frac{\alpha^2}{\langle\hat{B}\rangle^2}-\frac{\beta^2}{\langle\hat{A}\rangle^2}+\alpha^2\beta^2\right]. \end{aligned} \tag{5.54}$$

The interesting difference between this case and the sum of the errors case is that now the optimal probe state is dependent on the values of α and β, as the middle two terms have weightings of α^2 and β^2. Again the part we are interested in minimising can be reduced and given by,

$$\begin{aligned} e_{prod} &= \frac{1}{\langle\hat{A}\rangle^2\langle\hat{B}\rangle^2}-\frac{\alpha^2}{\langle\hat{B}\rangle^2}-\frac{\beta^2}{\langle\hat{A}\rangle^2}, \\ &= \frac{1}{\cos^2(\phi)\cos^2(2\theta-\phi)}-\frac{\beta^2}{\cos^2(\phi)}-\frac{\alpha^2}{\cos^2(2\theta-\phi)}. \end{aligned} \tag{5.55}$$

Differentiating with respect to ϕ we get the minimisation problem as,

$$\frac{2\tan\phi}{\cos^2\phi\cos^2(2\theta-\phi)}-\frac{2\tan(2\theta-\phi)}{\cos^2\phi\cos^2(2\theta-\phi)}-\frac{\beta^2\tan\phi}{\cos^2\phi}+\frac{\alpha^2\tan(2\theta-\phi)}{\cos^2(2\theta-\phi)}=0. \tag{5.56}$$

For an optimal measurement α and β are given by α_{opt} and β_{opt} respectively from equations (5.33) and (5.32). There are four choices for β_{opt} due to the plus and minus options

for each square root. The only unknown left is the probability p of which observable to be measured between $\hat{C}$ or $\hat{D}$. In figure 5.9 we chose $p = 0.7$ as that was the value used from the experimental setup in [61], where they use a beamsplitter with transmission probability p to choose between **c** and **d**.

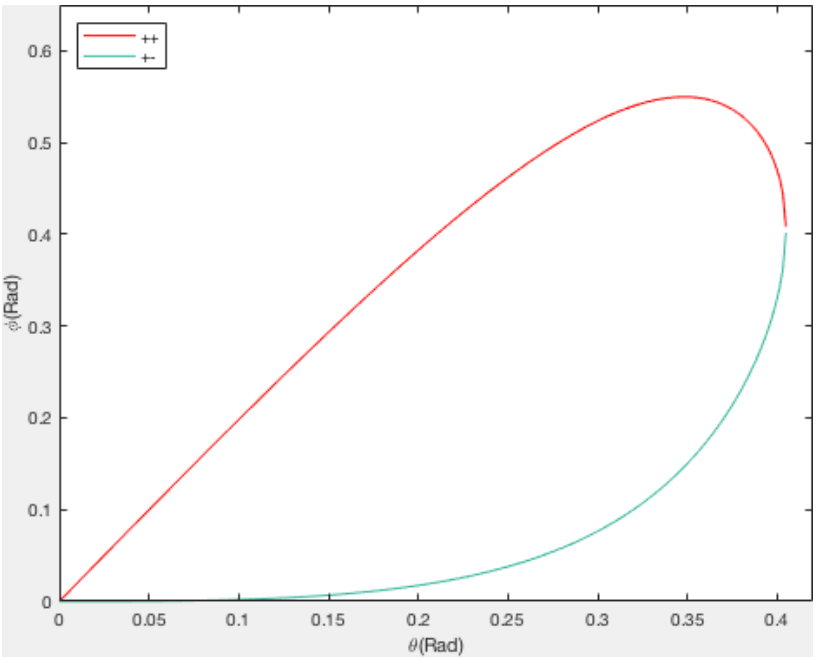


Figure 5.9: This figure shows the value of ϕ you would choose for each θ to minimise the error given by the term in equation (5.54) *with $p = 0.7$ and the sharpness of the measurements on $\hat{B}$ given by the β_{opt} solutions $\beta_{opt} = +\sqrt{+\sqrt{...}}$ and $\beta_{opt} = +\sqrt{-\sqrt{...}}$ from* (5.32).

All figures in this section are plotted from $\theta = 0$ to $\theta = \tan^{-1}((1-p)/p)$ as beyond this angle the measurement is not valid. From figure 5.9 we can see if we choose the β_{opt} solution $\beta_{opt} = +\sqrt{-\sqrt{...}}$, then for small values of θ the optimal probe state is close to $\phi = 0$. This means that when the two states **a** and **b** are close to each other then have a probe state near the eigenstate of **a** is ideal. As θ increases then the probe state increases more drastically and finishes at the point $\phi = \theta$ when θ is at the maximum angle determined by equation (5.35). For the $+-$ solution the optimum probe state stays close to $\phi = 2\theta$, which is the direction of **b**. We can understand this as looking at figure 5.10 we see that for small θ in for the $++$ solution $\beta \approx 1$ and therefore it makes sense to send the states along the **b** direction as this would make it a sharp measurement. Then as we reach the limiting angle $\alpha = \beta$ the probe state is halfway between the two. For the $+-$ solution α and β swap so for small θ, $\alpha \approx 1$, which explains why the probe state is near the **a** direction to begin with.

We can also alter the probability p, then we can see from figures 5.11a and 5.11b that the general shape stays the same but the curvature changes. For $p = 0.51$ the optimal probe state is either along the **a** direction or **b** direction depending on which solution

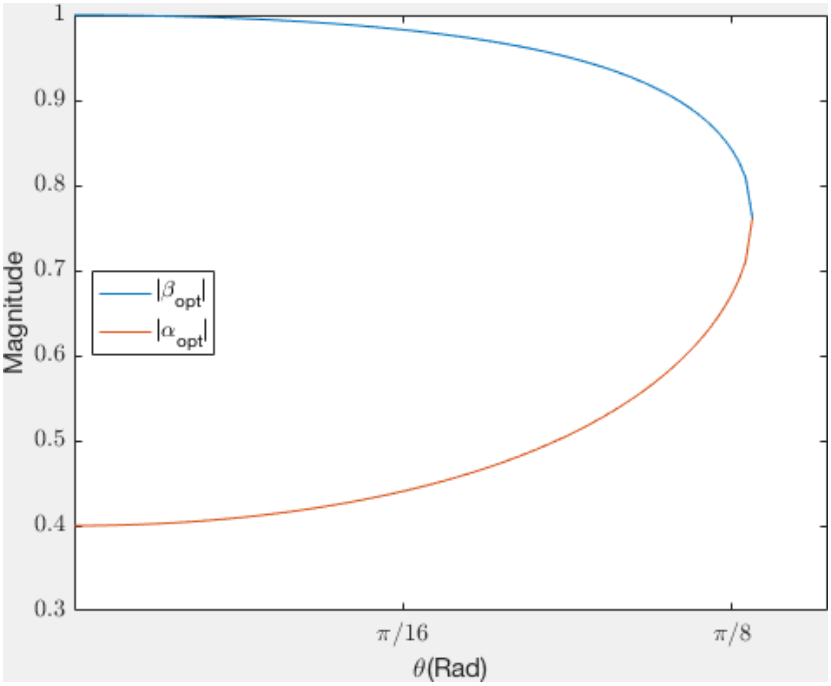
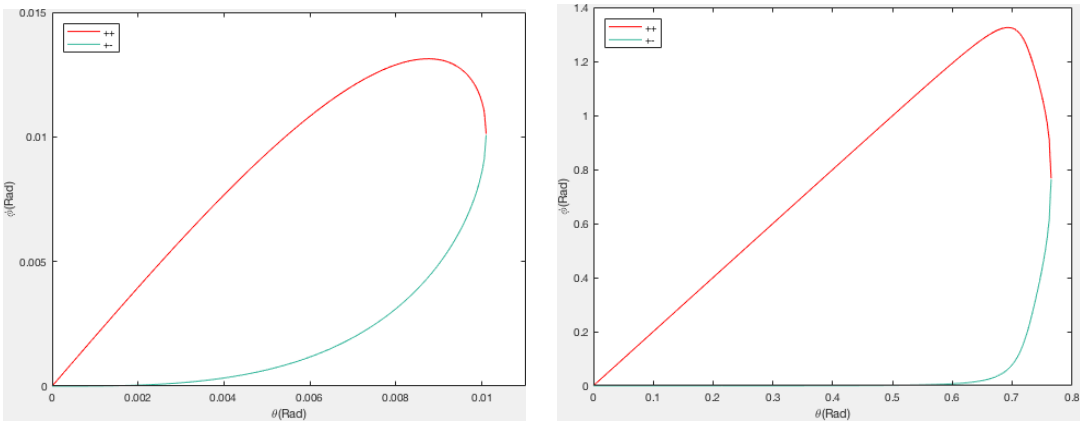


Figure 5.10: The magnitudes of β_{opt} and α_{opt} for $p = 0.7$ from (5.32) for the $++$ solution for θ from 0 to θ_{max} from equation (5.6).

of β is chosen. This is expected as for $p = 1/2$ we would have a measurement with $\alpha = 0, \beta = 1$ or $\alpha = 0, \beta = 1$, hence measurement along **a** or **b**.
In every case we have the optimum probe state for both solutions halfway between **a** and **b** for the limiting angle given in equation (5.35).



(a) $p = 0.99$ *(b)* $p = 0.51$

Figure 5.11: These figures shows the values of ϕ you would choose for each θ to minimise the error given by the term in equation (5.54) with $p = 0.99$ and $p = 0.51$. The sharpness of the measurements on $\hat{B}$ given by the β_{opt} solutions $\beta_{opt} = +\sqrt{+\sqrt{\dots}}$ and $\beta_{opt} = +\sqrt{-\sqrt{\dots}}$ from (5.32).

Figure 5.12 also helps show the difference between different p values. In this figure we have taken three p values of $p = 0.65, p = 0.7$ and $p = 0.75$. The valid angles for each

p vary so we have plotted up to the limiting angle for $p = 0.75$ as beyond this we would have at least one solution that is not valid.

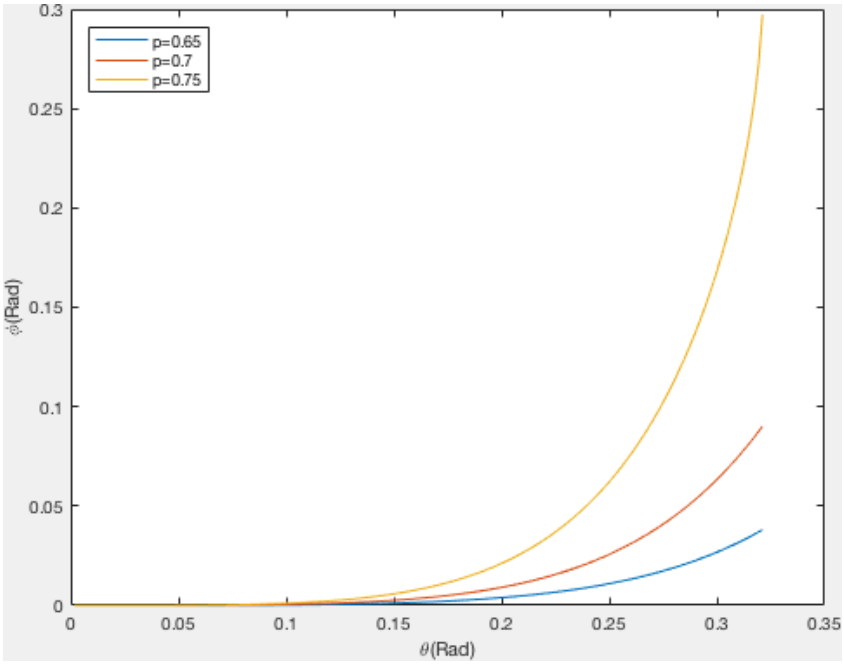


Figure 5.12: This figure shows the value of ϕ you would choose for each θ to minimise the error given by the term in (5.56)*. The sharpness of the measurement on $\hat{B}$ given by the β_{opt} solution $\beta_{opt} = +\sqrt{-\sqrt{...}}$ from* (5.32)*. Three different values of p have been used and we have plotted to the limiting angle of $p = 0.75$.*

5.4.3 Final Error Minimisation

With our collaborators in Bristol we also decided that it would be good to minimise the error in the estimation of,

$$\frac{(1-\alpha^2)(1-\beta^2)}{\alpha^2\beta^2} = \left(\frac{1}{\alpha^2} - 1\right)\left(\frac{1}{\beta^2} - 1\right). \tag{5.57}$$

This is the same function as the bound obtained in (5.6). To calculate the error we first find the error in the α part by taking the differential and obtaining,

$$\Delta_e\left(\frac{1}{\alpha^2} - 1\right) = -\frac{2}{\alpha^3}\Delta_e(\alpha). \tag{5.58}$$

The same can be done for β giving us

$$\Delta_e\left(\frac{1}{\beta^2} - 1\right) = -\frac{2}{\beta^3}\Delta_e(\beta). \tag{5.59}$$

Then using the product rule of differentiation we can see the total error will be given by

$$\Delta_e\left[\left(\frac{1}{\alpha^2} - 1\right)\left(\frac{1}{\beta^2} - 1\right)\right] = \left(\frac{-2}{\alpha^3}\right)\Delta_e(\alpha)\left(\frac{1}{\beta^2} - 1\right) + \left(\frac{1}{\alpha^2} - 1\right)\left(\frac{-2}{\beta^3}\right)\Delta_e(\beta), \tag{5.60}$$

where

$$\Delta_e(\alpha) = \sqrt{\frac{1}{N}\left(\frac{1}{\langle\hat{A}\rangle^2} - \alpha^2\right)} \qquad \text{and} \qquad \Delta_e(\beta) = \sqrt{\frac{1}{N}\left(\frac{1}{\langle\hat{B}\rangle^2} - \beta^2\right)}, \tag{5.61}$$

from equation (5.43). Again we can write $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ in terms of θ which we know and ϕ which is the variable describing the probe state. Then α and β are given by α_{opt} and β_{opt} respectively from equations (5.33) and (5.32) with some predetermined p. The analytic result from differentiating (5.60) is,

$$\frac{1}{N}\left[\frac{2\cos\phi\sin\phi(\beta^2-1)}{\alpha^3\beta^3\cos^3\phi\sqrt{1-\alpha^2\cos^2\phi}} - \frac{2\cos(2\theta-\phi)\sin(2\theta-\phi)(\alpha^2-1)}{\beta^3\alpha^3\cos^3(2\theta-\phi)\sqrt{1-\beta^2\cos^2(2\theta-\phi)}}\right]. \tag{5.62}$$

Figure 5.13 shows the ϕ to be chosen for each θ for a chosen p. Figure 5.13 and 5.8 are very similar both showing that for small θ for the $+-$ solution you send a state close to the **a** direction and for the $++$ solution you send the probe state close to the **b** direction. Even though the probe states for the product error and final error are very similar, from 5.14we can see they are in fact slightly different. Figure 5.14 shows the comparison of the optimum probe states for the $++$ solutions.

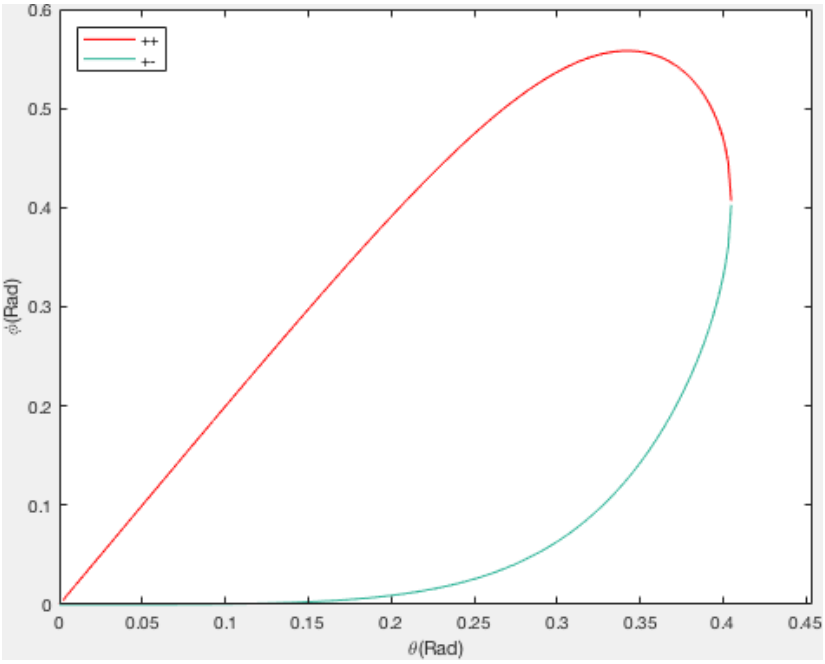


Figure 5.13: This figure shows the value of ϕ you would choose for each θ to minimise the error given by the term in 5.60 with $p = 0.7$ and the sharpness of the measurement on $\hat{B}$ given by the β_{opt} solutions $\beta_{opt} = +\sqrt{+\sqrt{\ldots}}$ and $\beta_{opt} = +\sqrt{-\sqrt{\ldots}}$ from equation (5.32).

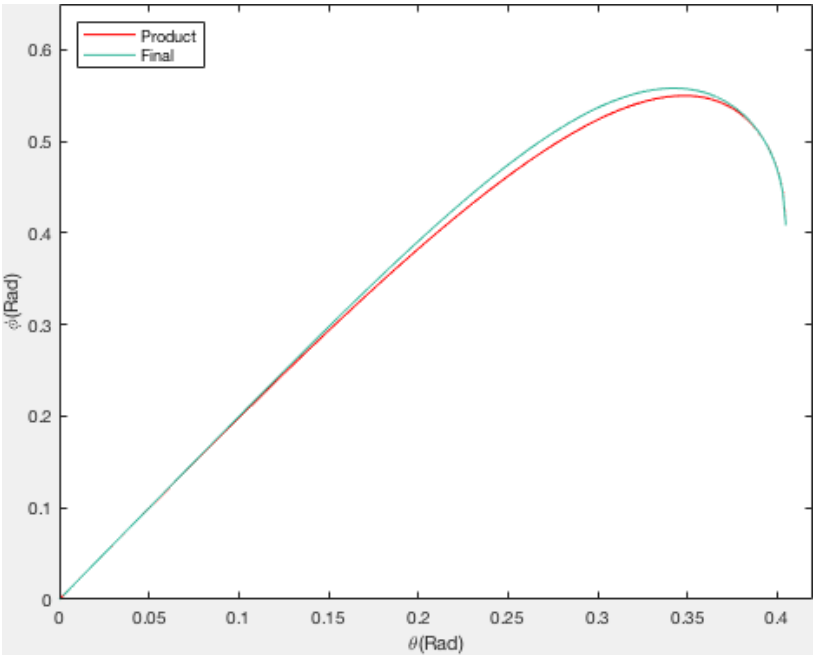


Figure 5.14: This figure shows the comparison of optimum probe states to minimise the product of errors and the final error minimisation for $p = 0.7$ and the $\beta_{opt} = +\sqrt{+\sqrt{...}}$ solution. The curves are similar but vary slightly with the value of ϕ lower for the product of error minimisation for some range of θ.

5.5 Conclusion

In this chapter we have looked at joint measurements and specifically how to minimise the error in the estimation of the sharpness of a measurement. This was done as a piece of work in co-operation with experimentalists and we managed to find the results required. As we based some of the work on the experiment we have looked at fixing p and varying θ to give us α and β values. Yet from a theoretical point of view it may be more intuitive to look at varying θ and also letting p change and thus being able to study all the optimal joint measurements for certain observables. This wouldn't restrict us to certain θ values for each p.

Chapter 6

Final Conclusion

In this thesis I have presented a selection of quantum elimination problems and then the numerical and analytical results for the optimal measurements involved to solve those problems. Semi-definite programming was used as a tool to obtain bounds on the success probabilities of elimination measurements. This form of convex optimisation turned out to be very useful and has obvious potential for many quantum applications as well as also being used for many purposes already, therefore it should definitely be considered for any current researchers looking for a numerical approach to obtain results. It would be interesting to investigate the duality gap that occurs in unambiguous measurements in more depth as it seems to be a fairly rare occurrence in quantum applications to not have strong duality.

The proof that a guaranteed optimal procedure to eliminate the highest average number of states is individual unambiguous measurements was a nice result for quantum information. Even though this may have been intuitive and even provable for minimum error measurements it was satisfying to have a solid proof of this for unambiguous measurements.

The applications from this work seem more to lie with the elimination measurements we obtained, especially eliminating multiple states. This can be seen by the quick QKD protocol we obtained, which has some different features to the more commonly used protocols. For example the two qubit systems make it possible to have a deterministic key production and not require such a thorough sifting system between Alice and Bob. As well as this the two out of four elimination seems like it could be used as an oblivious transfer protocol. I believe looking into the quantum communications applications of this measurement is the next obvious step to really make this work of practical relevance.

In chapter 4 we presented the process of decomposing a unitary into beamsplitter-like operations discovered by Reck et al. [58] so that a practical implementation could be

more easily visualised from the measurement unitary. This was applied to our two out of four measurement and with some level of optimisation we produced a relatively simple implementation. It would be interesting to collaborate with experimentalists to see if an experiment would be feasible and how we could make the implementation simpler.
I believe the decomposition has a more general application though and the ideal scenario would be to create some optimisation algorithm that took the initial measurement unitary and outputted the simplest experimental setup. The output would be dependent on some cost parameters that could be decided by each group depending on their desired style. We started forming an optimisation program but it had no certainties to produce ideal setup just *good* ones in general. I think this is an interesting area to look into as the reverse has been done with automated searches for experiments [69], so it seems a good idea to have an algorithm to produce an experimental setup.

In general I think elimination measurements could be investigated further with more general results found for eliminating say m out of n states. Then applications potentially lie in quantum communications or foundations as we have seen already in this thesis.

References

[1] Matthew F Pusey, Jonathan Barrett, and Terry Rudolph. On the reality of the quantum state. *Nature Physics*, 8(6):475–478, 2012.

[2] Stephen Barnett. *Quantum information*, volume 16. Oxford University Press, 2009.

[3] János A. Bergou. Tools for quantum information theory, 2010.

[4] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[5] John Von Neumann. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton university press, 2018.

[6] Mark Aronovich Naimark. On extremal spectral functions of a symmetric operator. In *Dokl. Akad. Nauk SSSR*, volume 54, page 9, 1946.

[7] Stephen M Barnett and Sarah Croke. Quantum state discrimination. *Advances in Optics and Photonics*, 1(2):238–278, 2009.

[8] Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. Conditions for compatibility of quantum-state assignments. *Physical Review A*, 66(6):062111, 2002.

[9] Somshubhro Bandyopadhyay, Rahul Jain, Jonathan Oppenheim, and Christopher Perry. Conclusive exclusion of quantum states. *Physical Review A*, 89(2):022336, 2014.

[10] Daniel KL Oi. Unlearning quantum information. *The European Physical Journal D*, 68(9):259, 2014.

[11] Igor D Ivanovic. How to differentiate between non-orthogonal states. *Physics Letters A*, 123(6):257–259, 1987.

[12] Dennis Dieks. Overlap and distinguishability of quantum states. *Physics Letters A*, 126(5-6):303–306, 1988.

[13] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1-2):19, 1988.

[14] Gregg Jaeger and Abner Shimony. Optimal distinction between two non-orthogonal quantum states. *Physics Letters A*, 197(2):83–87, 1995.

[15] Anthony Chefles. Unambiguous discrimination between linearly independent quantum states. *Physics Letters A*, 239(6):339–347, 1998.

[16] Anthony Chefles. Quantum state discrimination. *Contemporary Physics*, 41(6):401–424, 2000.

[17] János A Bergou, Ulrike Herzog, and Mark Hillery. 11 discrimination of quantum states. In *Quantum state estimation*, pages 417–465. Springer, 2004.

[18] Daniel Nigg, Thomas Monz, Philipp Schindler, Esteban A Martinez, Markus Hennrich, Rainer Blatt, Matthew F Pusey, Terry Rudolph, and Jonathan Barrett. Can different quantum state vectors correspond to the same physical state? an experimental test. *New Journal of Physics*, 18(1):013007, 2015.

[19] Martin Ringbauer, Ben Duffus, Cyril Branciard, Eric G Cavalcanti, Andrew G White, and Alessandro Fedrizzi. Measurements on the reality of the wavefunction. *Nature Physics*, 11(3):249, 2015.

[20] Kai-Yu Liao, Xin-Ding Zhang, Guang-Zhou Guo, Bao-Quan Ai, Hui Yan, and Shi-Liang Zhu. Experimental test of the no-go theorem for continuous ψ-epistemic models. *Scientific reports*, 6:26519, 2016.

[21] Jonathan Barrett, Eric G Cavalcanti, Raymond Lal, and Owen JE Maroney. No ψ-epistemic model can fully explain the indistinguishability of quantum states. *Physical review letters*, 112(25):250403, 2014.

[22] Matthew S Leifer. ψ-epistemic models are exponentially bad at explaining the distinguishability of quantum states. *Physical review letters*, 112(16):160404, 2014.

[23] Cyril Branciard. How ψ-epistemic models fail at explaining the indistinguishability of quantum states. *Physical review letters*, 113(2):020409, 2014.

[24] Janos Bergou, Edgar Feldman, and Mark Hillery. Extracting information from a qubit by multiple observers: Toward a theory of sequential state discrimination. *Physical review letters*, 111(10):100501, 2013.

[25] Lev B Levitin. Optimal quantum measurements for two pure and mixed states. In *Quantum Communications and Measurement*, pages 439–448. Springer, 1995.

[26] Robert J Collins, Ross J Donaldson, Vedran Dunjko, Petros Wallden, Patrick J Clarke, Erika Andersson, John Jeffers, and Gerald S Buller. Realization of quantum digital signatures without the requirement of quantum memory. *Physical review letters*, 113(4):040502, 2014.

[27] Christopher Perry, Rahul Jain, and Jonathan Oppenheim. Communication tasks with infinite quantum-classical separation. *Physical review letters*, 115(3):030504, 2015.

[28] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.

[29] Saul I Gass. Linear programming. *Encyclopedia of Statistical Sciences*, 6, 2004.

[30] Yonina C Eldar. A semidefinite programming approach to optimal unambiguous discrimination of quantum states. *IEEE Transactions on information theory*, 49(2):446–456, 2003.

[31] Andrew S Fletcher, Peter W Shor, and Moe Z Win. Optimum quantum error recovery using semidefinite programming. *Physical Review A*, 75(1):012338, 2007.

[32] Ryan Amiri and Juan Miguel Arrazola. Quantum money with nearly optimal error tolerance. *Physical Review A*, 95(6):062334, 2017.

[33] John Watrous. Lecture notes in theory of quantum information, October 2011.

[34] Jamie Sikora and Antonios Varvitsiotis. Semidefinite programming and quantum information.

[35] Carl Wilhelm Helstrom. *Quantum detection and estimation theory*. Academic press, 1976.

[36] Jamie Sikora. Semi-definite programming. QCrypt Lecture, 2017.

[37] Morton Slater. Lagrange multipliers revisited. In *Traces and Emergence of Nonlinear Programming*, pages 293–306. Springer, 2014.

[38] Michael Grant, Stephen Boyd, and Yinyu Ye. Cvx: Matlab software for disciplined convex programming, 2009.

[39] Erika Andersson and Mark Hillery. personal communication.

[40] Lev Vaidman, Yakir Aharonov, and David Z Albert. How to ascertain the values of sigma x, σ y, and σ z of a spin-1/2 particle. *Physical review letters*, 58(14):1385, 1987.

[41] Berthold-Georg Englert and Yakir Aharonov. The mean king's problem: prime degrees of freedom. *Physics Letters A*, 284(1):1–5, 2001.

[42] Berthold-Georg Englert, Christian Kurtsiefer, and Harald Weinfurter. Universal unitary gate for single-photon two-qubit states. *Physical Review A*, 63(3):032303, 2001.

[43] Almut Beige, Berthold-Georg Englert, Christian Kurtsiefer, and Harald Weinfurter. Secure communication with single-photon two-qubit states. *Journal of Physics A: Mathematical and General*, 35(28):L407, 2002.

[44] Sarah Croke. personal communication.

[45] Ronald L Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[46] Victor S Miller. Use of elliptic curves in cryptography. In *Conference on the theory and application of cryptographic techniques*, pages 417–426. Springer, 1985.

[47] Miklós Ajtai. Generating hard instances of lattice problems. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 99–108. ACM, 1996.

[48] Charles H Bennett and Gilles Brassard. Quantum cryptography: public key distribution and coin tossing. *Theor. Comput. Sci.*, 560(12):7–11, 2014.

[49] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Reviews of modern physics*, 81(3):1301, 2009.

[50] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature photonics*, 4(10):686, 2010.

[51] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-device-independent quantum key distribution. *Physical review letters*, 108(13):130503, 2012.

[52] William K Wootters and Wojciech H Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802, 1982.

[53] Peter W Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Physical review letters*, 85(2):441, 2000.

[54] Daniel Gottesman, H-K Lo, Norbert Lutkenhaus, and John Preskill. Security of quantum key distribution with imperfect devices. In *International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.*, page 136. IEEE, 2004.

[55] Charles H Bennett. Quantum cryptography using any two nonorthogonal states. *Physical review letters*, 68(21):3121, 1992.

[56] Hiroaki Sasaki, Ryutaroh Matsumoto, and Tomohiko Uyematsu. Key rate of the b92 quantum key distribution protocol with finite qubits. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 696–699. IEEE, 2015.

[57] Won-Young Hwang. Quantum key distribution with high loss: toward global secure communication. *Physical Review Letters*, 91(5):057901, 2003.

[58] Michael Reck, Anton Zeilinger, Herbert J Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Physical review letters*, 73(1):58, 1994.

[59] Thorlabs. Thorlabs optical elements.

[60] Jonathan Crickmore, Jonathan Frazer, Scott Shaw, and Pieter Kok. Effect of component variations on the gate fidelity in linear optical networks. *Physical Review A*, 94(2):022326, 2016.

[61] Adetunmise C Dada, Will McCutcheon, Erika Andersson, Jonathan Crickmore, Ittoop Puthoor, Brian D Gerardot, Alex McMillan, John Rarity, and Ruth Oulton. Optimal simultaneous measurements of incompatible observables of a single photon. *Optica*, 6(3):257–263, 2019.

[62] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, and Umesh Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. *arXiv preprint quant-ph/9804043*, 1998.

[63] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, and Umesh Vazirani. Dense quantum coding and quantum finite automata. *Journal of the ACM (JACM)*, 49(4):496–511, 2002.

[64] Andris Ambainis, Debbie Leung, Laura Mancinska, and Maris Ozols. Quantum random access codes with shared randomness. *arXiv preprint arXiv:0810.2937*, 2008.

[65] Thomas Brougham and Erika Andersson. Estimating the expectation values of spin-1/2 observables with finite resources. *Physical Review A*, 76(5):052313, 2007.

[66] E Arthurs and MS Goodman. Quantum correlations: A generalized heisenberg uncertainty relation. *Physical review letters*, 60(24):2447, 1988.

[67] Paul Busch. Unsharp reality and joint measurements for spin observables. *Physical Review D*, 33(8):2253, 1986.

[68] Erika Andersson, Stephen M Barnett, and Alain Aspect. Joint measurements of spin, operational locality, and uncertainty. *Physical Review A*, 72(4):042104, 2005.

[69] Mario Krenn, Mehul Malik, Robert Fickler, Radek Lapkiewicz, and Anton Zeilinger. Automated search for new quantum experiments. *Physical review letters*, 116(9):090405, 2016.